(19) Weltorganisation für geistiges Eigentum Internationales Büro





(43) Internationales Veröffentlichungsdatum 29. Juli 2004 (29.07.2004)

PCT

(10) Internationale Veröffentlichungsnummer WO 2004/063359 A2

- (51) Internationale Patentklassifikation⁷: C12N 1/15, 15/80, C12P 23/00, A23J 1/00, 3/00, A23L 1/28, 1/275
- (21) Internationales Aktenzeichen: PCT/EP2004/000099
- (22) Internationales Anmeldedatum:

9. Januar 2004 (09.01.2004)

(25) Einreichungssprache:

Deutsch

(26) Veröffentlichungssprache:

Deutsch

(30) Angaben zur Priorität:

103 00 649.4 103 41 271.9 9. Januar 2003 (09.01.2003) DE 8. September 2003 (08.09.2003) DE

(71) Anmelder (für alle Bestimmungsstaaten mit Ausnahme von US): BASF AKTIENGESELLSCHAFT [DE/DE]; 67056 Ludwigshafen (DE).

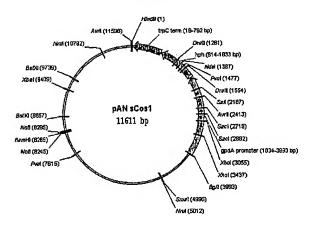
- (72) Erfinder; und
- (75) Erfinder/Anmelder (nur für US): MATUSCHEK, Markus [DE/DE]; Karolinenstr. 5, 69469 Weinheim (DE). KLEIN, Daniela [DE/DE]; M 7, 2, 68161 Mannheim (DE). HEINEKAMP, Thorsten [DE/DE]; Alte Ziegelei 1B, 30419 Hannover (DE). SCHMIDT, Andre [DE/DE]; Magdeburger Str. 11, 31832 Springe (DE). BRAKHAGE, Axel [DE/DE]; Schneiderberg 58, 30167 Hannover (DE). ACHATZ, Brigitte [DE/DE]; Windeckstr. 26, 68163 Mannheim (DE).
- (74) Anwalt: FITZNER, Uwe; Lintorfer Str. 10, 40878 Ratingen (DE).
- (81) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare nationale Schutzrechtsart): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES,

[Fortsetzung auf der nächsten Seite]

(54) Title: METHOD FOR PRODUCING CAROTENOIDS OR THEIR PRECURSORS USING GENETICALLY MODIFIED ORGANISMS OF THE BLAKESLEA GENUS, CAROTENOIDS OR THEIR PRECURSORS PRODUCED BY SAID METHOD AND USE THEREOF

(54) Bezeichnung: VERFAHREN ZUR HERSTELLUNG VON CAROTINOIDEN ODER DEREN VORSTUFEN MITTELS GENTECHNISCH VERÄNDERTER ORGANISMEN DER GATTUNG BLAKESLEA, MIT DEM VERFAHREN HERGESTELLTE CAROTINOIDE ODER DEREN VORSTUFEN UND DEREN VERWENDUNG

Vektor pANsCos1



(57) Abstract: The invention relates to a method for producing carotenoids or their precursors using genetically modified organisms of the *Blakeslea* genus. Said method comprises the following steps (i) transformation of at least one of the cells, (ii) optional homokaryotic conversion of the cells obtained in step (i) to produce cells, in which one or more genetic characteristics of the nucleii are all modified in an identical manner and said modification manifests itself in the cells, (iii) selection and reproduction of the genetically modified cell or cells, (iv) cultivation of the genetically modified cells, (v) preparation of the carotenoids produced by the genetically modified cells or the carotenoid precursor produced by said genetically modified cells. The invention also relates to carotenoids or their precursors produced according to said method and to the use thereof.

WO 2004/063359

WO 2004/063359 A2

FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(84) Bestimmungsstaaten (soweit nicht anders angegeben, für jede verfügbare regionale Schutzrechtsart): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), eurasisches (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Veröffentlicht:

 ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts

Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen. WO 2004/063359 PCT/EP2004/000099

Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderter Organismen der Gattung Blakeslea, mit dem Verfahren hergestellte Carotinoide oder deren Vorstufen und deren Verwendung

5

10

15

Die Erfindung betrifft ein Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderten Organismen der Gattung Blakeslea, mit dem Verfahren hergestellte Carotinoide oder deren Vorstufen und deren Verwendung und Bereitstellung, besonders als hochreine Carotinoide, als Nahrungsmittel, enthaltend Carotinoide-produzierende Organismen und mindestens ein Carotinoid, insbesondere Tierfuttermittel, Tierfutterergänzungsmittel und Nahrungsergänzungsmittel, sowie die Verwendung der aus dem Verfahren erhältlichen Carotinoide zur Herstellung von kosmetischen, pharmazeutischen, dermatologischen Zubereitungen, Nahrungsmitteln oder Nahrungsergänzungsmitteln.

Blakeslea trispora ist als Produktionsorganismus für β-Carotin (Ciegler, 1965, Adv Appl Microbiol. 7:1) und Lycopin bekannt (EP 1201762, EP 1184464, WO 03/038064).

20

Von Blakeslea trispora sind bisher verschiedene DNA-Sequenzen bekannt, insbesondere die DNA-Sequenz, die für die Gene der Carotinoid-biosynthese von Geranylgeranylpyrophosphat bis β -Carotin codiert (WO 03/027293).

25 -

Insbesondere aufgrund der hohen Produktivität, die mit Blakeslea in der Produktion von Lycopin und β-Carotin erreicht werden, bietet sich dieser Organismus zur fermentativen Herstellung von Carotinoiden an.

30 Es ist auch von Interesse die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Her-

10

15

20

25

stellung weiterer Carotinoide, wie z. B. Xanthophylle zu ermöglichen, die von Blakeslea bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Carotinoide werden Futtermitteln, Nahrungsmitteln, Nahrungsergänzungsmitteln, Kosmetika und Arzneimitteln zugesetzt. Die Carotinoide dienen vor allem als Pigmente zur Färbung. Daneben werden die antioxidative Wirkung der Carotinoide und andere Eigenschaften dieser Substanzen genutzt. Man unterteilt die Carotinoide in die reinen Kohlenwasserstoffe, die Carotine und die sauerstoffhaltigen Kohlenwasserstoffe, die Xanthophylle. Xanthophylle wie Canthaxanthin und Astaxanthin werden beispielsweise zur Pigmentierung von Hühnereiern und Fischen eingesetzt (Britton et al. 1998, Carotinoids, Vol 3, Biosynthesis and Metabolism). Die Carotine β-Carotin und Lycopin werden vor allem in der Humanernährung eingesetzt. β -Carotin wird beispielsweise als Getränkefarbstoff verwendet. Lycopin hat eine krankheitsvorbeugende Wirkung (Argwal und Rao, 2000, CMAJ 163:739-744; Rao und Argwal 1999, Nutrition Research 19:305-323). Die farblose Carotinoidvorstufe Phytoen kommt vor allem für Anwendungen als Antioxidans in kosmetischen, pharmazeutischen oder dermatologischen Zubereitungen in Frage.

Der überwiegende Teil der Carotinoide und deren Vorstufen, die als Zusatzstoffe für die oben genannten Anwendungen eingesetzt werden, wird durch chemische Synthese hergestellt. Die chemische Synthese ist technisch sehr aufwendig und verursacht hohe Herstellkosten. Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden und deren Vorstufen können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder neue Carotinoide auf Basis der bekannten

WO 2004/063359 PCT/EP2004/000099

würde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten.

Hierzu ist eine gentechnische, d. h. gezielte genetische Veränderung von Blakeslea erforderlich. Insbesondere, wenn Xanthophylle produziert werden sollen, da diese Verbindungen natürlicherweise vom Wildtyp der Blakeslea nicht synthetisiert werden.

• 5

15

20

30

- Z. B. zur Herstellung von Phytoen mittels Fermentation von Blakeslestrispora sind bisher zwei Methoden bekannt:
 - (i) Durch zufallsabhängige Mutagenese mit chemischen Agenzien wie MNNG können Mutanten erzeugt werden, in denen Phytoen nicht zu Lycopin und somit nicht weiter zu ß-Carotin umgesetzt werden kann (Mehta und Cerdá-Olmedo, 1995, Appl. Microbiol. Biotechnol. 42:836-838).
 - (ii) Durch Zugabe von Inhibitoren des Enzyms Phytoendesaturase wie z.B. Diphenylamin und Zimtalkohol kann die weitere Umsetzung von Phytoen blockiert werden, so dass es sich anreichert (Cerdá-Olmedo, 1989, In: E. Vandamme, ed. Biotechnoiogy of vitamin, growth factorand pigment production. London: Eisevier Applied Science, S. 27-42).

Die genannten Methoden zur Herstellung von Phytoen mit Blakeslea 25 trispora weisen jedoch eine Reihe von Nachteilen auf.

Die zufallsabhängie Mutagenese betrifft in der Regel nicht nur die Gene der Carotinoidbiosynthese zur weiteren Umsetzung von Phytoen, sondern auch weitere wichtige Gene. Daher sind Wachstum und Syntheseleistung der Mutanten oft beeinträchtigt. Die Erzeugung z. B. von Phytoenüberpro-

25

duzenten durch zufallsabhängige Mutagenese von Lycopinüberproduzenten oder ß-Carotinüberproduzenten ist daher entweder nicht oder nur mit großem experimentellem Aufwand zu erreichen. Die Zugabe von Inhibitoren verursacht eine Erhöhung der Produktionskosten und gegebenenfalls eine Verunreinigung des Produktes. Daneben kann das Zellwachstum durch den Inhibitor beeinträchtigt werden, so dass die Produktion von Carotinoiden oder deren Vorstufen, insbesondere Phytoen eingeschränkt wird.

- Durch eine gentechnische Veränderung könnten die oben genannten Nachteile der zufallsabhängigen Mutagenese und der Inhibitorzugabe vermieden werden.
- Allerdings sind bisher keine Methoden zur gentechnischen, d. h. gezielten gentechnischen Veränderung von Blakeslea, insbesondere Blakeslea trispora bekannt.

Als Methode zur Herstellung von gentechnisch veränderten Pilzen wurde in einigen Fällen die Agrobacterium-vermittelte Transformation erfolgreich eingesetzt. So sind z. B. folgende Organismen durch Agrobakterien transformiert worden: Saccharomyces cerevisiae (Bundock et al., 1995, EMBO Journal, 14:3206–3214), Aspergillus awamori, Aspergillus nidulans, Aspergillus niger, Colletotrichum gloeosporioides, Fusarium solani pisi, Neurospora crassa, Trichoderma reesei, Pleurotus ostreatus, Fusarium graminearum (van der Toorren et al., 1997, EP 870835), Agraricus bisporus, Fusarium venenatum (de Groot et al., 1998, Nature Biotechnol. 16:839–842), Mycosphaerella graminicola (Zwiers et al. 2001, Curr. Genet. 39:388–393), Glarea lozoyensis (Zhang et al., 2003, Mol. Gen. Genet.

15

20

nomics 268:645–655), Mucor miehei (Monfort et al. 2003, FEMS Microbiology Lett. 244:101 – 106).

Von Interesse ist besonders eine homologe Rekombination, bei der zwischen der einzuführenden DNA und der Zell-DNA möglichst viele Sequenzhomologien bestehen, so dass eine ortsspezifische Einführung bzw. Ausschaltung von genetischer Information im Genom des Empfängerorganismus möglich ist. Andernfalls wird die Spender-DNA durch illegitime bzw. nicht-homologe Rekombination ins Genom des Empfängerorganismus integriert, was nicht ortsspezifisch erfolgt.

Eine durch Agrobacterium vermittelte Transformation und anschließende homologe Rekombination der transferierten DNA wurde bisher bei folgenden Organismen nachgewiesen: Aspergillus awamori (Gouka et al. 1999, Nature Biotech 17:598-601), Glarea lozoyensis (Zhang et al., 2003, Mol. Gen. Genomics 268:645-655), Mycosphaerella graminicola ((Zwiers et al. 2001, Curr. Genet. 39:388-393).

Als weitere Methode zur Transformation von Pilzen ist die Elektroporation bekannt. Die integrative Transformation von Hefe durch Elektroporation wurde von Hill, Nucl. Acids. Res. 17:8011 gezeigt. Für filamentöse Pilze wurde die Transformation durch Chakaborty und Kapoor beschrieben (1990, Nucl. Acids. Res. 18:6737).

Eine "biolistische" Methode, d.h. die Übertragung von DNA durch Beschuss von Zellen mit DNA-beladenen Partikeln wurde beispielsweise für Trichoderma harzianum und Gliocladium virens beschrieben (Lorito et al. 1993, Curr. Genet. 24:349–356).

WO 2004/063359

25

30

Diese Methoden konnten bisher jedoch nicht erfolgreich zur gezielten genetischen Veränderung von Blakeslea und insbesondere Blakeslea trispora eingesetzt werden.

- 5 Eine besondere Schwierigkeit bei der Herstellung von gentechnisch veränderten Blakeslea und Blakeslea trispora, ist die Tatsache, dass deren Zellen in allen Stadien des sexuellen und des vegetativen Zellzyklus mehrkernig sind. In Sporen von Blakeslea trispora Stamm NRRL2456 und NRRL2457 wurden z. B. im Durchschnitt 4,5 Kerne pro Spore nachgewiesen (Metha und Cerdá-Olmedo, 1995, Appl. Microbiol. Biotechnol. 42:836–838). Dies hat zur Folge, dass die gentechnische Veränderung in aller Regel nur in einem oder wenigen Kernen vorliegt, die Zellen also heterokaryotisch sind.
- Sollen die genetisch veränderten Blakeslea, insbesondere Blakeslea trispora zur Produktion eingesetzt werden, so ist es insbesondere bei einer Gendeletion wichtig, dass in den Produktionsstämmen die gentechnische Veränderung in allen Kernen vorliegt, so dass eine stabile und hohe Syntheseleistung ohne Nebenprodukte möglich wird. Die Stämme müssen folglich in Bezug auf die gentechnische Veränderung homokaryotisch sein.

Lediglich für Phycomyces blakesleeanus ist ein Verfahren beschrieben worden, um homokaryotische Zellen zu erzeugen (Roncero et al., 1984, Mutat. Res. 125:195). Durch Zugabe des mutagenen Agens MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) werden nach dem dort beschriebenen Verfahren Kerne in den Zellen eliminiert, so dass statistisch eine gewisse Anzahl von Zellen mit nur noch einem funktionellem Kern vorliegt. Die Zellen werden dann einer Selektion unterzogen, in der nur einkernige Zellen mit einem rezessiven Selektionsmarker zu einem Mycel auswachsen können. Die Nachkommen dieser selektierten Zellen sind mehrkernig und homokaryotisch. Ein rezessiver Selektionsmarker für Phycomyces blakes-

WO 2004/063359 PCT/EP2004/000099

7

leanus ist z. B. dar. dar⁺-Stämme nehmen das toxische Riboflavin-Analog 5-Carbon-5-deazariboflavin auf; dar⁻-Stämme dagegen nicht (Delbrück et al. 1979, Genetics 92:27). Die Selektion von rezessiven Mutanten erfolgt durch Zugabe von 5-Carbon-5-deazariboflavin (DARF).

5

Allerdings ist dieses Verfahren nicht für Blakeslea, insbesondere Blakeslea trispora bekannt und insbesondere nicht mit im Zusammenhang mit einer Transformation oder der Produktion von Carotinoiden oder deren Vorstufen beschrieben worden.

10

15

Auch die Isolierung aus natürlichen Quellen wird durchgeführt. Beispielsweise ist es für die Gewinnung von Phytoen bekannt, ein Gemisch aus Carotinoiden, Vitamin E und anderen Komponenten, welches auch Phytoen enthält, aus Tomaten, Karotten oder Palmöl usw. zu extrahieren. Problematisch ist hierbei die Trennung der einzelnen Carotinoide voneinander. So ist beispielsweise das Phytoen nach diesem Verfahren nicht in reiner Form erhältlich. Insbesondere ist die natürlich vorkommende Menge der Carotinoiden in den Pflanzen gering.

20

25

30

Fermentative Verfahren sind demgegenüber technisch verhältnismäßig einfach und basieren auf kostengünstigen Einsatzstoffen. Fermentative Verfahren zur Herstellung von Carotinoiden können dann wirtschaftlich attraktiv und wettbewerbsfähig zur chemischen Synthese sein, wenn die Produktivität der bisherigen fermentativen Verfahren gesteigert würde oder neue Carotinoide auf Basis der bekannten Produktionsorganismen hergestellt werden könnten. Problematisch bei der fermentativen Herstellung von Carotinoiden sind allerdings die Aufarbeitungsverfahren, die nur geringe Mengen an hochreinen Carotinoiden bereitstellen. Zudem sind dafür meist aufwendige Vielschritt-Prozesse ggf. unter Verwendung großer Lösungsmittelmengen erforderlich. So fallen große Mengen Abfall an oder es

10

15

muss ein hoher Aufwand zur Wiederverwertung (Recycling) betrieben werden.

Die Produktion von Carotinoiden durch verschiedene Mikroorganismen ist an sich bekannt. So ist z. B. in der WO 00/13654 A2 offenbart, ein Gemisch aus Phytoen und Phytofluen aus Algen der Art Dunaliella sp. zu extrahieren. Auch nach diesem Verfahren ist das Phytoen nicht in reiner Form erhältlich und muss von den anderen Produkten getrennt werden. Zudem handelt es sich um gentechnisch unveränderte Algen, deren Biosynthese mittels eines hinzugefügten Inhibitors beeinflusst werden muss.

Blakeslea trispora als Produktionsorganismus für β -Carotin ist auch aus der WO 98/03480 A1. bekannt. Hier werden β -Carotin Kristalle aus Biomasse von Blakeslea trispora mittels Extraktion erhalten. Allerdings müssen in dem beschriebenen Verfahren große Mengen unterschiedlicher Lösungsmittel eingesetzt werden, um Kristalle mit hoher Reinheit durch mehrere Extraktions- und Waschschritte zu erhalten. Auch sind die erhaltenen Mengen β -Carotin bezogen auf die eingesetzte Menge Biomasse klein.

Aus der WO 01/83437 A1 ist ein Verfahren zur Extraktion von Astaxanthin aus Hefe bekannt, bei dem die Kulturbrühe zur Sterilisation und zum Zell-aufschluss mit Mikrowellenstrahlung behandelt wird. Der Zellaufschluss mittels Mikrowellenstrahlung ist danach nötig, um Astaxanthin aus Hefe zu gewinnen, ohne es dabei zu zerstören. Anschließend soll Astaxanthin mittels Methanol, Ethanol oder Aceton oder deren Mischungen extrahiert werden. Hierzu sind allerdings große Mengen Lösungsmittel (5 bis 20 Teile Lösungsmittel auf 1 Teil Suspension) und ein langer Zeitraum (24h) erforderlich. Zudem sind keine Reinheiten des Astaxanthins angegeben und die erhaltenen Mengen sind klein. Versuche der Anmelderin und andere Veröffentlichungen bestätigen jedoch, dass eine Extraktion mittels Methanol oder Ethanol nicht durchführbar ist.

Aus der WO 98/50574 ist ebenfalls die Isolierung von Carotinoid Kristallen aus Biomasse von Mikroorganismen bekannt, wobei hiernach im Gegensatz zur WO 01/83437 A1 Methanol, Ethanol, Aceton nur zum Entfernen von Lipiden aus der Biomasse d. h. .zum Waschen verwendet werden kann. Als Lösungsmittel zur Extraktion von Carotinoiden wird demnach Ethylacetat, Hexan oder ein Öl verwendet. Anschließend sind mehrere Reinigungs- und Waschschritte mit großen Mengen Ethanol und Wasser nötig, wobei lediglich eine Reinheit von 93,3 % bei einer Ausbeute von 35 % erreicht wird.

10

15

20

25

30

Die WO 03/038064 A2 beschreibt die fermentative Produktion von Lycopin durch Co-Kultivierung von mutiertem Blakeslea trispora Paarungstyp (–) und Blakeslea trispora Paarungstyp (+), die ohne Zusatz von Inhibitoren der Carotinoid Biosynthese Lycopin herstellen. Die Erzeugung der zur Fermentation eingesetzten Mutante wird durch unselektive chemische Mutation und anschließendes Screening vorgenommen. Die Aufarbeitung der Kulturbrühe erfolgt mittels Zellaufschluss und anschließender Reinigung mit unterschiedlichen wässrigen Medien mit verschiedenem Salzgehalt und pH-Wert und mit Wasser nicht mischbaren organischen Lösungsmitteln wie Ethylacetat, Hexan und 1- Butanol zur Entfernung von Lipiden. Alternativ ist eine Extraktion mittels großer Mengen Ethylacetat beschrieben. Angaben zur Reinheit fehlen. Da Ethylacetat und Hexan Lösungsmittel für Lycopin sind, ist davon auszugehen, dass ein Teil des Lycopins herausgewaschen und so die theoretische mögliche Ausbeute verringert wird.

Auch aus der WO 01/55100 A1 ist die Isolierung von Carotinoiden allgemein bzw. β-Carotin im speziellen aus der Biomasse durch Anwendung mehrerer Wasch- und Reinigungsschritte auf die aufgeschlossene Biomasse ohne Extraktion mittels Lösungsmittel beschrieben. Hierzu wird

aufgeschlossene Biomasse von Blakeslea trispora mit Wasser, Lauge, Säure, Butanol und Ethanol gewaschen, so daß eine große Zahl unterschiedlicher Lösungsmittel und wässriger Medien verwendet werden muss. Die Reinheit des erhaltenen β -Carotins beträgt 96 – 98 %. Angaben zur Ausbeute fehlen jedoch.

Die WO 97/36996 A2 beschreibt allgemein eine Verfahren zur Isolierung von Substanzen (u. a. Carotinoide) aus Mikroorganismen, wobei die Substanzen aus der Biomasse mittels Fest/Flüssig-Extraktion isoliert werden. Ein Zellaufschluß soll hierbei nicht nötig sein, jedoch muss die Biomasse zunächst durch Extrusion in eine granulierte, poröse Gestalt gebracht werden. Wie nur Carotinoide isoliert werden können und wie deren Reinheit bzw. Ausbeute ist, ist nicht angegeben. Der Rückstand der Extrusion kann anschließend als Futtermittelzusatz verwendet werden.

15

20

25

30

5

10

In allen oben beschriebenen Verfahren müssen große Mengen Lösungsmittel zur Extraktion eingesetzt werden, um die isolierte Menge an Carotinoid durch vollständige Extraktion zu erhöhen, und/oder große Mengen wässriger Medien zur Reinigung und zum Waschen eingesetzt werden. Dies bedingt hohe Kosten und aufwendige Maßnahmen zur Wiederverwendung bzw. ggf. Abfälle.

Zudem werden die nahrhafte Kulturbrühe und die darin enthaltene Biomasse nach Extraktion bzw. Isolierung der Carotinoide als Abfall behandelt. Die oben angegebenen Verfahren haben neben diesen vordergründigen Nachteilen einen entscheidenden weiteren Nachteil. Es ist nämlich danach notwendig, die Carotinoide den Nahrungsmitteln nachträglich zuzusetzen, d. h. sie sind nicht Bestandteil der Nahrungsmittel an sich bzw. nicht in ausreichender Menge. Von großem Vorteil wäre daher, wenn der Gehalt an Carotinoiden in den Nahrungsmitteln bereits durch die eigentlichen Nahrungsmittel selbst gedeckt würde.

WO 2004/063359 PCT/EP2004/000099

11

Es ist ebenfalls nötig die Produktivitäten der bisher natürlicherweise produzierten Carotine und deren Vorstufen weiter zu steigern und die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin zu ermöglichen, die von den Wildtypen der Mikroorganismen bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

Aufgabe der Erfindung ist es gentechnisch veränderte Zellen von Blakeslea-Stämmen, insbesondere Blakeslea trispora bereitzustellen, die Carotinoide oder deren Vorstufen, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin produzieren.
Zudem soll das Verfahren die Steigerung der Carotinoid-Produktivität der
veränderten Zellen gegenüber den korrespondierenden Wildtypen erlauben. Ferner soll das Verfahren die Erzeugung neuer Zellen oder aus ihnen
bestehendes Mycel erlauben, die sich für die Verwendung zur Herstellung
von Carotinoiden oder deren Vorstufen eignen, die bisher nicht in wirtschaftlich interessanten Mengen aus den natürlich vorkommenden Pilzen
gewinnbar waren, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin. Das Verfahren soll dabei
eine gentechnische Veränderung von Blakeslea-Stämmen, insbesondere
Blakeslea trispora möglich machen und die Herstellung homokaryotischer
gentechnisch veränderter Produktions-Stämme erlauben.

Des weiteren soll das Verfahren die Herstellung weiterer Carotinoide, wie z. B. Xanthophylle, insbesondere Astaxanthin oder Zeaxanthin und Phytoen oder Bixin ermöglichen, die von den Wildtypen der Mikroorganismen bisher nicht oder nur in sehr geringem Maße gebildet und isoliert werden können.

5

10

15

20

10

15

20

25

Ferner ist es Aufgabe der vorliegenden Erfindung ein Verfahren zur Herstellung von Carotinoiden aus gentechnisch veränderte Zellen von Blakeslea-Stämmen, insbesondere Blakeslea trispora, zur Verfügung zu stellen, welches den Einsatz geringerer Lösungsmittelmengen erlaubt und im wesentlichen ohne Abfälle auskommt und zudem eine hohe Reinheit und höhere Ausbeuten erlaubt.

In diesem Zusammenhang soll ein möglichst großer Anteil der im Fermenter vorliegenden Nährstoffe, sowohl Carotinoide als auch weitere sich in den Mikroorganismen befindende, verwertet werden.

Somit ist es auch Aufgabe der vorliegenden Erfindung ein Verfahren zur Herstellung eines Carotinoid-haltigen Nahrungsmittels bereitzustellen, wobei das Nahrungsmittel selbst den Bedarf an Carotinoiden ohne Zusätze deckt. Insbesondere soll der Nährstoffgehalt der nach dem Verfahren erhältlichen Nahrungsmittel gegenüber den bisher erhältlichen Nahrungsmitteln zumindest gleichwertig sein. Ferner soll das Verfahren die effiziente Verwertung der produzierten Carotinoide ermöglichen.

Diese Aufgabe wird durch ein Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderten Organismen der Gattung Blakeslea gelöst, umfassend

- (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
- (iii) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen,
- (iv) Kultivierung der gentechnisch veränderten Zellen,

- Bereitstellung des von den gentechnisch veränderten Zellen (v) produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe.
- Mit der erfindungsgemäßen Methode ist es möglich, Blakeslea gezielt und 5 stabil genetisch zu verändern, um so Mycel aus Zellen mit einheitlichen Kernen zu gewinnen, das Carotinoide oder deren Vorstufen, insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin und Phytoen oder Bixin produziert. Vorzugsweise handelt es sich um Zellen von Pilzen der Art Blakeslea trispora. Die produzierten Carotinoiden oder de-10 ren Vorstufen sind dabei im wesentlichen frei von Verunreinigungen erhältlich und es können hohe Konzentrationen der Carotinoiden oder deren Vorstufen im Kulturmedium erzielt werden.

- Unter Transformation wird die Übertragung einer genetischen Information in den Organismus, insbesondere Pilz verstanden. Darunter sollen alle dem Fachmann bekannten Möglichkeiten zur Einschleusung der Information, insbesondere DNA fallen, z. B. Beschuss mit DNA-beladenen Partikeln, Transformation mittels Protoplasten, Mikroinjektion von DNA, Elektroporation, Konjugation oder Transformation kompetenter Zellen, Chemi-20 kalien oder Agrobakterien vermittelte Transformation. Als genetische Information werden ein Genabschnitt, ein Gen oder mehrere Gene verstanden. Die genetische Information kann z. B. mit Hilfe eines Vectors oder als freie Nukleinsäure (z. B. DNA, RNA) und auf sonstige Weise in die Zellen eingebracht und entweder durch Rekombination ins Wirtsgenom einge-25 baut oder in freier Form in der Zelle vorliegen. Besonders bevorzugt ist hierbei die homologe Rekombination.
- Bevorzugte Transformationsmethode ist die Agrobacterium tumefaciensvermittelte Transformation. Hierzu wird zunächst die zu transferierende 30

Spender-DNA in einen Vektor eingefügt, der (i) flankierend zu der zu transferierenden DNA die T-DNA-Enden trägt, der (ii) einen Selektionsmarker enthält und der (iii) ggf. Promotoren und Terminatoren für die Genexpression der Spender-DNA aufweist. Dieser Vektor wird in einen Agrobacterium-tumefaciens-Stamm übertragen, der ein Ti-Plasmid mit den vir-Genen enthält. vir-Gene sind für den DNA-Transfer in Blakeslea verantwortlich. Mit diesem Zwei-Vektor-System wird die DNA von Agrobacterium in Blakeslea übertragen. Hierzu werden die Agrobakterien zunächst in Gegenwart von Acetosyringone inkubiert. Acetosyringone induziert die vir-Gene. Anschließend werden Sporen von Blakeslea trispora zusammen mit den induzierten Zellen von Agrobacterium tumefaciens auf Acetosyringone-haltigem Medium inkubiert und dann auf Medium übertragen, das eine Selektion der Transformanten, d.h. der gentechnisch veränderten Stämme von Blakeslea ermöglicht.

15

20

25

30

10

Der Begriff Vector wird in der vorliegenden Anmeldung als eine Bezeichnung für ein DNA-Molekül verwendet, das zum Einschleusen und ggf. zur Vermehrung von Fremd-DNA in eine Zelle dient (siehe auch "Vector" in Römpp Lexikon Chemie – CDROM Version 2.0, Stuttgart/New York: Georg Thieme Verlag 1999). In der vorliegenden Anmeldung sollen unter dem begriff "Vector" auch Plasmide, Cosmide usw. verstanden werden, die dem gleichen Zweck dienen.

Unter Expression wird in der vorliegenden Anmeldung die Übertragung einer genetischen Information ausgehend von DNA oder RNA in ein Gen-Produkt (hier vorzugsweise Enzyme zur Herstellung von Carotinoiden und insbesondere Xanthophylle, besonders bevorzugt Astaxanthin oder Zea-xanthin und Phytoen oder Bixin) verstanden und soll auch den Begriff der Überexpression beinhalten, womit eine verstärkte Expression gemeint ist,

so dass ein bereits in der nicht transformierten Zelle (Wildtyp) hergestell-

tes Genprodukt verstärkt produziert wird oder einen großen Teil des gesamten Gehaltes der Zelle ausmacht.

Unter gentechnische Veränderung soll die Einschleusung genetischer Information in einen Empfängerorganismus, so dass diese stabil exprimiert und bei der Zellteilung weitergegeben wird, verstanden werden. In diesem Zusammenhang ist die Homokaryotisierung, die Herstellung von Zellen, die nur einheitliche Kerne enthalten, d. h. Kerne mit gleichem genetischem Informationsgehalt.

10

5

Diese Homokaryotisierung ist nur notwendig, wenn die durch Transformation eingeführte genetische Information rezessiv vorliegt, d. h. nicht zur Ausprägung gelangt. Führt die Transformation aber zu einem dominanten Vorliegen der genetischen Information, d. h. wird sie ausgeprägt, so ist eine Homokaryotisierung nicht unbedingt nötig.

15

20

Vorzugsweise wird zur Homokaryotisierung eine Selektion der einkernigen Sporen durchgeführt. Von Natur aus ist ein geringer Anteil der Sporen von Blakeslea trispora einkernig, so dass sich diese ggf. nach spezifischer Markierung z. B. Färbung der Zellkerne aussortieren lassen. Dies wird bevorzugterweise mittels FACS (Fluorescence Activated Cell Sorting) anhand der geringeren Fluoreszenz der einkernigen Zellen durchgeführt.

25

30

Alternativ kann zur Homokaryotisierung zunächst eine Kernreduktion durchgeführt werden. Hierzu kann ein mutagenes Agens eingesetzt werden, wobei es sich insbesondere um N-Methyl-N'-nitro-nitrosoguanidin (MNNG) handelt. Auch die Verwendung von energiereichen Strahlen, wie UV- oder Röntgen-Strahlen zur Kernreduktion ist möglich. Anschließend kann zur Selektion auf das FACS Verfahren oder rezessive Selektionsmarker zurückgegriffen werden.

Unter Selektion wird die Auswahl von Zellen verstanden, deren Kerne dieselbe genetische Information beinhalten, d. h. Zellen die die gleichen Eigenschaften aufweisen, wie Resistenzen oder die Herstellung bzw. vermehrte Herstellung eines Produktes. In der Selektion werden neben der FACS Methode bevorzugt 5-Carbon-5-deazariboflavin (DARF) und Hygromycin (hyg) oder 5'-Fluororotat (FOA) und Uracil eingesetzt.

Der in der Transformation (i) eingesetzte Vector kann derart gestaltet sein, dass die im Vector enthaltene genetische Information in das Genom mindestens einer Zelle integriert wird. Dabei kann genetische Information in der Zelle ausgeschaltet werden. Dies kann direkt, d. h. durch eine Deletion erfolgen. Es ist aber auch möglich, daß der in der Transformation (i) eingesetzte Vector derart ausgestaltet ist, dass die im Vector enthaltene genetische Information in der Zelle exprimiert wird, d. h. genetische Information eingefügt wird, die im korrespondierenden Wildtyp nicht vorhanden ist oder die durch die Transformation verstärkt bzw. überexprimiert wird und deren Produkt das Gen ausschaltet. Die eingeführte genetische Information kann aber auch indirekt eine genetische Information in der Zelle ausschalten, z. B. durch Produktion eines Inhibitors.

20

15

5

10

Der eingesetzte Vector enthält genetische Informationen oder Teile der genetischen Information zur Herstellung von Carotinoiden oder deren Vorstufen, insbesondere Carotinen oder Xanthophyllen oder deren Vorstufen. Der eingesetzte Vector enthält vorzugsweise genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β -Cryptoxanthin, β -Carotin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, Lutein, Bixin oder Phytoen. Ganz besonders bevorzugt enthält der Vector Informationen zur Herstellung von Bixin, Phytoen, Canthaxanthin, Astaxanthin oder Zeaxanthin.

25

Der Vector kann beliebige genetische Informationen zur genetischen Veränderungen von Organismen der Gattung Blakeslea enthalten.

Unter "genetischer Information" werden vorzugsweise Nukleinsäuren verstanden, deren Einbringung in den Organismus der Gattung Blakeslea zu einer genetischen Veränderung in Organismen der Gattung Blakeslea, also beispielsweise zu einer Verursachung, Erhöhung oder Reduzierung von Enzymaktivitäten im Vergleich zum Ausgangsorganismus führen.

10 Der Vector kann beispielsweise genetische Information zur Herstellung lipophiler Substanzen enthalten wie z.B. Carotinoide und deren Vorstufen, Phospholipide, Triacylglyceride, Steroide, Wachse, fettlösliche Vitamine, Provitamine und Cofaktoren oder genetische Information zur Herstellung hydrophiler Substanzen wie z.B. Eiweiße, Aminosäuren, Nukleotide und wasserlösliche Vitaminen, Provitamine und Cofaktoren.

Bevorzugterweise enthält der eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder Xanthophyllen oder deren Vorstufen.

20

5

Bevorzugterweise enthält der Vektor genetische Information, die eine Lokalisierung der Carotinoidbiosynthese-Enzyme in dem Zellkompartiment bewirkt, in dem die Carotinoidbiosynthese stattfindet.

Besonders bevorzugt sind genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β-Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3- und 3'-Hydroxyechinenon, Lycopin, Lutein, β-Carotin, Phytoen und/oder Phytofluen. Ganz besonders bevorzugt sind genetische Informationen zur Herstellung von Phytoen, Bixin, Lycopin, Zeaxanthin, Canthaxanthin und/oder Astaxanthin.

Entsprechend werden in einer bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die über eine erhöhte Syntheserate für Zwischenprodukte der Carotinoidbiosynthese verfügen und folglich eine erhöhte Produktivität für Endprodukte der Carotinoidbiosynthese aufweisen. Zur Erhöhung der Syntheserate für Zwischenprodukte der Carotinoidbiosynthese werden insbesondere die Aktivitäten der Enzyme 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase (HMG-CoA-Reduktase), Isopentenylpyrophosphat-Isomerase und Geranylpyrophosphatsynthase gesteigert.

10

20

25

30

5

Entsprechend werden in einer besonders bevorzugten Variante der Erfindung Organismen hergestellt und kultiviert, die gegenüber dem Wildtypeine erhöhte HMG-CoA-Reduktase-Aktivität aufweisen.

Unter HMG-CoA-Reduktase—Aktivität wird die Enzymaktivität einer HMG-CoA-Reduktase (3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A-Reduktase) verstanden.

Unter einer HMG-CoA-Reduktase wird ein Protein verstanden, das die enzymatische Aktivität aufweist 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A in Mevalonat umzuwandeln.

Dementsprechend wird unter HMG-CoA-Reduktase—Aktivität die in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. gebildete Menge Mevalonat verstanden.

Bei einer erhöhten HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein HMG-CoA-Reduktase die umgesetzte Menge 3-Hydroxy-3-Methyl-Glutaryl-Coenzym-A bzw. die gebildete Menge Mevalonat erhöht.

Vorzugsweise beträgt diese Erhöhung der HMG-CoA-Reduktase-Aktivität mindestens 5%, weiter bevorzugt mindestens 20%, weiter bevorzugt mindestens 50%, weiter bevorzugt mindestens 100%, besonders bevorzugt mindestens 300%, noch bevorzugter mindestens 500%, insbesondere mindestens 600% der HMG-CoA-Reduktase-Aktivität des Wildtyps.

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der HMG-CoA-Reduktase-Aktivität gegenüber dem Wildtyp durch eine Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase.

10

15

5

In einer besonders bevorzugten Ausführungsform des erfindungsgemäßen Verfahrens erfolgt die Erhöhung der Genexpression einer Nukleinsäure codierend eine HMG-CoA-Reduktase indem man ein Nukleinsäurekonstrukt, enthaltend eine Nukleinsäure codierend eine HMG-CoA-Reduktase in den Organismus einbringt, deren Expression in dem Organismus, verglichen mit dem Wildtyp, einer reduzierten Regulation unterliegt.

Unter einer reduzierten Regulation verglichen mit dem Wildtyp, wird eine im Vergleich zum vorstehend definierten Wildtyp verringerte, vorzugsweise keine Regulation auf Expressions- oder Proteinebene verstanden.

20 keine Regulatio

Die reduzierte Regulation kann vorzugsweise durch einen im Nukleinsäurekonstrukt mit der kodierenden Sequenz funktionell verknüpften Promotor erreicht werden, der in dem Organismus verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

Beispielsweise unterliegen die Promotoren ptef1 aus Blakeslea trispora und pgpdA aus Aspergillus nidulans nur einer reduzierten Regulation und sind daher insbesondere als Promotoren bevorzugt.

25

Diese Promotoren zeigen eine annähernd konstitutive Expression in Blakeslea trispora, so dass die transkriptionelle Regulation nicht mehr über die Intermediate der Carotinoidbiosynthese abläuft.

Die reduzierte Regulation kann in einer weiteren bevorzugten Ausführungsform dadurch erreicht werden, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt.

10

15

20

25

Besonders bevorzugt ist die Verwendung einer Nukleinsäure, die nur den katalytischen Bereich der HMG-CoA-Reduktase kodiert (trunkierte (t-)HMG-CoA-Reduktase). Die für die Regulation verantwortliche Membran-Domäne fehlt. Die verwendete Nukleinsäure unterliegt somit einer reduzierten Regulation und führt zu einer Erhöhung der Genexpression der HMG-CoA-Reduktase.

In einer besonders bevorzugten Ausführungsform bringt man Nukleinsäuren in Blakeslea trispora ein, welche die Sequenz SEQ ID. NO. 75 enthalten.

Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Homologievergleiche der Sequenzen aus Datenbanken mit der SEQ ID. NO. 75 leicht auffinden.

Weitere Beispiele für HMG-CoA-Reduktasen und damit auch für die auf den katalytischen Bereich reduzierten t-HMG-CoA-Reduktasen bzw. die kodierenden Gene lassen sich weiterhin beispielsweise ausgehend von

der Sequenz SEQ ID. NO. 75 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

In einer besonders bevorzugten Ausführungsform wird die reduzierte Regulation dadurch erreicht, dass man als Nukleinsäure codierend eine HMG-CoA-Reduktase eine Nukleinsäure verwendet, deren Expression in dem Organismus, verglichen mit der Organismus eigenen, orthologen Nukleinsäure, einer reduzierten Regulation unterliegt und einen Promotor verwendet, der in dem Organismus, verglichen mit dem Wildtyp-Promoter einer reduzierten Regulation unterliegt.

Entsprechend wird in einer bevorzugten Variante der Erfindung durch die Transformation die Genexpression der Phytoendesaturase ausgeschaltet, so dass das von den Organismen produzierte Phytoen gewonnen werden kann. Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Phytoendesaturase, insbesondere carB aus Blakeslea trispora mit der SEQ ID NO: 69.

20

25

15

Entsprechend wird in einer bevorzugten Variante der Erfindung durch Transformation die Genexpression der Lycopincyclase ausgeschaltet, so dass das von den Organismen produzierte Lycopin gewonnen werden kann. Der in der Transformation eingesetzte Vektor umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für ein Fragment des Gens der Lycopincyclase, insbesondere carR aus Blakeslea trispora.

30 In einer bevorzugten Ausführungsform werden die Organismen der Gattung Blakeslea beispielsweise dadurch in die Lage versetzt Xanthophylle, wie beispielsweise Canthaxanthin, Zeaxanthin oder Astaxanthin herzustellen, Bixin oder Phytoen, indem in den genetisch veränderten Organismen der Gattung Blakeslea im Vergleich zum Wildtyp eine Hydroxylase-Aktivität und/oder Ketolase-Aktivität verursacht wird.

5

Der in der Transformation (i) eingesetzte Vector enthält also in einer weiteren, bevorzugten Variante der Erfindung genetische Informationen, die nach Expression eine Ketolase- und/oder Hydroxylase-Aktivität entfalten, so dass die Organismen Zeaxanthin oder Astaxanthin produzieren.

10

Unter Ketolase-Aktivität wird die Enzymaktivität einer Ketolase verstanden.

l In

Unter einer Ketolase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -lonon-Ring von Carotinoiden eine Keto-Gruppe einzuführen.

Insbesondere wird unter einer Ketolase ein Protein verstanden, das die enzymatische Aktivität aufweist, β -Carotin in Canthaxanthin umzuwandeln.

20

15

Dementsprechend wird unter Ketolase–Aktivität die in einer bestimmten Zeit durch das Protein Ketolase umgesetzte Menge β -Carotin bzw. gebildete Menge Canthaxanthin verstanden.

25 L

Unter dem Begriff "Wildtyp" wird erfindungsgemäß der entsprechende nicht genetisch veränderte Ausgangsorganismus der Gattung Blakesleaa verstanden.

30

Je nach Zusammenhang kann unter dem Begriff "Organismus" der Ausgangsorganismus (Wildtyp) der Gattung Blakesleaa oder ein erfindungs-

WO 2004/063359

gemäßer, genetisch veränderter Organismus der Gattung Blakesleaa oder beides verstanden werden.

23

PCT/EP2004/000099

Vorzugsweise wird unter "Wildtyp" für die Verursachung der Ketolase-Aktivität und für die Verursachung der Hydroxylase-Aktivität jeweils eine Referenz Organismus verstanden.

Dieser Referenzorganismus der Gattung Blakeslea ist Blakeslea trispora ATCC 14271 oder ATCC 14272, die sich lediglich im Paarungstyp unterscheiden.

Die Bestimmung der Ketolase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen der Gattung Blakesleaa und in Wildtyp- bzw. Referenzorganismen erfolgt vorzugsweise unter folgenden Bedingungen:

15

20

25

30

10

Die Bestimmung der Ketolase-Aktivität in Organismen der Gattung Blakeslea erfolgt in Anlehnung an die Methode von Frazer et al., (J. Biol. Chem. 272(10): 6128-6135, 1997). Die Ketolase-Aktivität in Extrakten wird mit den Substraten beta-Carotin und Canthaxanthin in Gegenwart von Lipid (Sojalecithin) und Detergens (Natriumcholat) bestimmt. Substrat/Produkt-Verhältnisse aus den Ketolase-Assays werden mittels HPLC ermittelt.

Der erfindungsgemäße genetisch veränderte Organismus der Gattung Blakesleaa weist in dieser, bevorzugten Ausführungsform im Vergleich zum genetisch nicht veränderten Wildtyp eine Ketolase-Aktivität auf und ist somit vorzugsweise in der Lage, transgen eine Ketolase zu exprimieren.

In einer weiter bevorzugten Ausführungsform erfolgt die Verursachung der Ketolase-Aktivität in den Organismen der Gattung Blakesleaa durch Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase.

WO 2004/063359 PCT/EP2004/000099

In dieser bevorzugten Ausführungsform erfolgt die Verursachung der Genexpression einer Nukleinsäure kodierend eine Ketolase vorzugsweise durch Einbringen von Nukleinsäuren, die Ketolasen kodieren in die Ausgangsorganismus der Gattung Blakesleaa.

Dazu kann prinzipiell jedes Ketolase-Gen, also jede Nukleinsäuren die eine Ketolase codiert verwendet werden.

Alle in der Beschreibung erwähnten Nukleinsäuren können beispielsweise eine RNA-, DNA- oder cDNA-Sequenz sein.

Bei genomischen Ketolase-Sequenzen aus eukaryontischen Quellen, die Introns enthalten, sind für den Fall das der Wirtsorganismus der Gattung Blakesleaa nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechenden Ketolase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

20 Beispiele für Nukleinsäuren, kodierend eine Ketolase und die entsprechenden Ketolasen, die im erfindungsgemäßen Verfahren verwendet werden können sind beispielsweise Sequenzen aus:

Haematoccus pluvialis, insbesondere aus Haematoccus pluvialis Flotow em. Wille (Accession NO: X86782; Nukleinsäure: SEQ ID NO: 11, Protein SEQ ID NO: 12),

Haematoccus pluvialis, NIES-144 (Accession NO: D45881; Nukleinsäure: SEQ ID NO: 13, Protein SEQ ID NO: 14),

5

15

Agrobacterium aurantiacum (Accession NO: D58420; Nukleinsäure: SEQ ID NO: 15, Protein SEQ ID NO: 16),

Alicaligenes spec. (Accession NO: D58422; Nukleinsäure: SEQ ID NO: 17, Protein SEQ ID NO: 18),

Paracoccus marcusii (Accession NO: Y15112; Nukleinsäure: SEQ ID NO: 19, Protein SEQ ID NO: 20).

Synechocystis sp. Strain PC6803 (Accession NO: NP442491; Nukleinsäure: SEQ ID NO: 21, Protein SEQ ID NO: 22).

Bradyrhizobium sp. (Accession NO: AF218415; Nukleinsäure: SEQ ID NO: 23, Protein SEQ ID NO: 24).

15

25

30

Nostoc sp. Strain PCC7120 (Accession NO: AP003592, BAB74888; Nukleinsäure: SEQ ID NO: 25, Protein SEQ ID NO: 26),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No.

NZ_AABC01000195, Basenpaar 55,604 bis 55,392 (SEQ ID NO: 27); Protein: Acc.-No. ZP_00111258 (SEQ ID NO: 28) (als putatives Protein annotiert),

Nostoc punctiforme ATTC 29133, Nukleinsäure: Acc.-No.

NZ_AABC01000196, Basenpaar 140,571 bis 139,810 (SEQ ID NO: 29),

Protein: (SEQ ID NO: 30) (nicht annotiert),

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene, die im erfindungsgemäßen Verfahren verwendet werden können, lassen sich beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, durch Identitätsvergleiche der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit den vorstehend beschriebenen Sequenzen und insbesondere mit den Sequenzen SEQ ID NO: 12, 26 und/oder 33 leicht auffinden.

Weitere natürliche Beispiele für Ketolasen und Ketolase-Gene lassen sich weiterhin ausgehend von den vorstehend beschriebenen Nukleinsäuresequenzen, insbesondere ausgehend von den Sequenzen SEQ ID NO: 12,
26 und/oder 30 aus verschiedenen Organismen, deren genomische Sequenz nicht bekannt ist, durch Hybridisierungstechniken in an sich bekannter Weise leicht auffinden.

Die Hybridisierung kann unter moderaten (geringe Stringenz) oder vorzugsweise unter stringenten (hohe Stringenz) Bedingungen erfolgen.

Solche Hybridisierungsbedingungen sind beispielsweise bei Sambrook, J., Fritsch, E.F., Maniatis, T., in: Molecular Cloning (A Laboratory Manual), 2. Auflage, Cold Spring Harbor Laboratory Press, 1989, Seiten 9.31-9.57 oder in Current Protocols in Molecular Biology, John Wiley & Sons, N.Y. (1989), 6.3.1-6.3.6 beschrieben.

20

Beispielhaft können die Bedingungen während des Waschschrittes ausgewählt sein aus dem Bereich von Bedingungen begrenzt von solchen mit geringer Stringenz (mit 2X SSC bei 50_C) und solchen mit hoher Stringenz (mit 0.2X SSC bei 50_C, bevorzugt bei 65_C) (20X SSC: 0,3 M Natriumcitrat, 3 M Natriumchlorid, pH 7.0).

Darüberhinaus kann die Temperatur während des Waschschrittes von moderaten Bedingungen bei Raumtemperatur, 22°C, bis zu stringenten Bedingungen bei 65°C angehoben werden.

25

Beide Parameter, Salzkonzentration und Temperatur, können gleichzeitig variiert werden, auch kann einer der beiden Parameter konstant gehalten und nur der andere variiert werden. Während der Hybridisierung können auch denaturierende Agenzien wie zum Beispiel Formamid oder SDS eingesetzt werden. In Gegenwart von 50% Formamid wird die Hybridisierung bevorzugt bei 42°C ausgeführt.

Einige beispielhafte Bedingungen für Hybridisierung und Waschschritt sind infolge gegeben:

10

- (1) Hybridiserungsbedingungen mit zum Beispiel
- ¹ (i) 4X SSC bei 65°C, oder
- (ii) 6X SSC bei 45°C, oder
- (iii) 6X SSC bei 68°C, 100 mg/ml denaturierter Fischsperma-DNA, oder
- 15 (iv) 6X SSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA bei 68°C, oder
 - (v) 6XSSC, 0.5 % SDS, 100 mg/ml denaturierte, fragmentierte Lachssperma-DNA, 50 % Formamid bei 42°C, oder
 - (vi) 50 % Formamid, 4X SSC bei 42°C, oder
- (vii) 50 % (vol/vol) Formamid, 0.1 % Rinderserumalbumin, 0.1 % Ficoll,
 0.1 % Polyvinylpyrrolidon, 50 mM Natriumphosphatpuffer pH 6.5, 750 mM
 NaCl, 75 mM Natriumcitrat bei 42°C, oder
 - (viii) 2X oder 4X SSC bei 50°C (moderate Bedingungen), oder
- (ix) 30 bis 40 % Formamid, 2X oder 4X SSC bei 42°C (moderate Bedingungen).
 - (2) Waschschritte für jeweils 10 Minuten mit zum Beispiel
 - (i) 0.015 M NaCl/0.0015 M Natriumcitrat/0.1 % SDS bei 50°C, oder
 - (ii) 0.1X SSC bei 65°C, oder
- 30 (iii) 0.1X SSC, 0.5 % SDS bei 68°C, oder
 - (iv) 0.1X SSC, 0.5 % SDS, 50 % Formamid bei 42°C, oder

25

30

- (v) 0.2X SSC, 0.1 % SDS bei 42°C, oder
- (vi) 2X SSC bei 65°C (moderate Bedingungen)...

In einer bevorzugten Ausführungsform der erfindungsgemäßen genetisch veränderten Organismen der Gattung Blakeslea bringt man Nukleinsäuren ein, die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 12 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30%, 40%, 50%, 60%, bevorzugt mindestens 70%, 80%, besonders bevorzugt mindestens 90%, insbesondere 91%, 92%; 93%, 94%, 95%, 96%, 97%, 98% oder 99% auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die wie vorstehend beschrieben durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 12 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 26 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30%, 40%, 50%, 60%, bevorzugt mindestens 70%, 80%, besonders bevorzugt mindestens 90%, insbesondere 91%, 92%; 93%, 94%, 95%, 96%, 97%, 98% oder 99% auf Aminosäureebene mit der Sequenz SEQ ID NO: 26 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 26 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

5

10

15

30

In einer weiteren, bevorzugten Ausführungsform der erfindungsgemäßen Verfahren bringt man Nukleinsäuren ein die ein Protein kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 30 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 20 %, vorzugsweise mindestens 30 %, 40 %, 50 %, bevorzugt mindestens 60 %, 70 %, bevorzugter mindestens 80 %, 85 % besonders bevorzugt mindestens 90 %, insbesondere 91 %, 92 %, 93 %, 94 %, 95 %, 96 %, 97 %, 98 %, 99 % auf Aminosäureebene mit der Sequenz SEQ ID NO 30 und die enzymatische Eigenschaft einer Ketolase aufweist.

Dabei kann es sich um eine natürliche Ketolase-Sequenz handeln, die, wie vorstehend beschrieben, durch Identitätsvergleich der Sequenzen aus anderen Organismen gefunden werden kann oder um eine künstliche Ketolase-Sequenz die ausgehend von der Sequenz SEQ ID NO: 30 durch künstliche Variation, beispielsweise durch Substitution, Insertion oder Deletion von Aminosäuren abgewandelt wurde.

Unter dem Begriff "Substitution" ist in der Beschreibung der Austausch einer oder mehrerer Aminosäuren durch eine oder mehrere Aminosäuren zu verstehen. Bevorzugt werden sog. konservative Austausche durchgeführt, bei denen die ersetzte Aminosäure eine ähnliche Eigenschaft hat

wie die ursprüngliche Aminosäure, beispielsweise Austausch von Glu durch Asp, Gln durch Asn, Val durch Ile, Leu durch Ile, Ser durch Thr.

Deletion ist das Ersetzen einer Aminosäure durch eine direkte Bindung. Bevorzugte Positionen für Deletionen sind die Termini des Polypeptides und die Verknüpfungen zwischen den einzelnen Proteindomänen.

Insertionen sind Einfügungen von Aminosäuren in die Polypeptidkette, wobei formal eine direkte Bindung durch ein oder mehrere Aminosäuren wird. ersetzt

Unter Identität zwischen zwei Proteinen wird die Identität der Aminosäuren über die jeweils gesamte Proteinlänge verstanden, insbesondere die Identität die durch Vergleich mit Hilfe der Lasergene Software der Firma DNASTAR, inc. Madison, Wisconsin (USA) unter Anwendung der Clustal Methode (Higgins DG, Sharp PM. Fast and sensitive multiple sequence alignments on a microcomputer. Comput Appl. Biosci. 1989 Apr;5(2):151-1) unter Einstellung folgender Parameter berechnet wird:

Multiple alignment parameter: 20

10

15

10 Gap penalty 10 Gap length penalty Pairwise alignment parameter: 1

K-tuple 3 Gap penalty 25 5 Window 5 Diagonals saved

Unter einem Protein, das eine Identität von mindestens 20% auf Aminosäureebene mit der Sequenz SEQ ID NO: 12 oder 26 oder 30 aufweist, 30 wird dementsprechend ein Protein verstanden, das bei einem Vergleich seiner Sequenz mit der Sequenz SEQ ID NO: 12 oder 26 oder 30, insbesondere nach obigen Programmlogarithmus mit obigem Parametersatz eine Identität von mindestens 20 %, bevorzugt 30%, 40%, 50%, besonders bevorzugt 60%, 70%, 80%, insbesondere 85%, 90, 95% aufweist.

5

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

0

Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Blakesleaaspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene von Organismen der Gattung Blakesleaa leicht ermitteln.

15

In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 11 in die Organismus der Gattung ein.

_ _

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 25 in die Organismus der Gattung ein.

20

In einer weiteren, besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ ID NO: 29 in die Organismus der Gattung ein.

25

30

Alle vorstehend erwähnten Ketolase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender, komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, S. 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

Der in der Transformation (i) eingesetzte Vector umfasst daher in einer Ausführungsform der Erfindung bevorzugterweise eine Sequenz codierend für eine Ketolase, insbesondere der Ketolase Nostoc punctiforme aus mit der SEQ ID NO: 72.

Unter Hydroxylase-Aktivität die Enzymaktivität einer Hydroxylase verstanden.

15

25

10

5

Unter einer Hydroxylase wird ein Protein verstanden, das die enzymatische Aktivität aufweist, am, gegebenenfalls substituierten, β -lonon-Ring von Carotinoiden eine Hydroxy-Gruppe einzuführen.

Insbesondere wird unter einer Hydroxylase ein Protein verstanden, das die enzymatische Aktivität aufweist, β-Carotin in Zeaxanthin oder Cantaxanthin in Astaxanthin umzuwandeln.

Dementsprechend wird unter Hydroxyase–Aktivität die in einer bestimmten Zeit durch das Protein Hydroxylase umgesetzte Menge β-Carotin oder Cantaxanthin bzw. gebildete Menge Zeaxanthin oder Astaxanthin verstanden.

Bei einer erhöhten Hydroxylase-Aktivität gegenüber dem Wildtyp wird somit im Vergleich zum Wildtyp in einer bestimmten Zeit durch das Protein

WO 2004/063359 PCT/EP2004/000099

33

Hydroxylase die umgesetzte Menge β-Carotin oder Cantaxantin bzw. die gebildete Menge Zeaxanthin oder Astaxanthin erhöht.

Vorzugsweise beträgt diese Erhöhung der Hydroxylase–Aktivität mindestens 5 %, weiter bevorzugt mindestens 20 %, weiter bevorzugt mindestens 50 %, weiter bevorzugt mindestens 100 %, bevorzugter mindestens 300 %, noch bevorzugter mindestens 500 %, insbesondere mindestens 600 % der Hydroxylase–Aktivität des Wildtyps.

Die Bestimmung der Hydroxylase-Aktivität in erfindungsgemäßen genetisch veränderten Organismen und in Wildtyp- bzw. ReferenzOrganismen erfolgt vorzugsweise unter folgenden Bedingungen:

Die Aktivität der Hydroxylase wird nach Bouvier et al. (Biochim. Biophys. Acta 1391 (1998), 320-328) in vitro bestimmt. Es wird zu einer bestimmten Menge an Organismenextrakt Ferredoxin, Ferredoxin-NADP Oxidoreductase, Katalase, NADPH sowie beta-Carotin mit Mono- und Digalaktosylglyzeriden zugegeben.

20 Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase–Aktivität unter folgenden Bedingungen nach Bouvier, Keller, d'Harlingue und Camara (Xanthophyll biosynthesis: molecular and functional characterization of carotenoid hydroxylases from pepper fruits (Capsicum annuum L.; Biochim. Biophys. Acta 1391 (1998), 320-328):

25

30

15

Der in-vitro Assay wird in einem Volumen von 0.250 ml Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6), 0.025 mg Ferredoxin von Spinat, 0.5 Einheiten Ferredoxin-NADP+ Oxidoreduktase von Spinat, 0.25 mM NADPH, 0.010 mg beta-Carotin (in 0.1 mg Tween 80 emulgiert), 0.05 mM einer Mischung von Mono- und Digalaktosylglyzeriden (1:1), 1 Einheit Katalyse, 200 Mono- und Digalaktosylglyzeriden, (1:1),

0.2 mg Rinderserumalbumin und Organismenextrakt in unterschiedlichem Volumen. Die Reaktionsmischung wird 2 Stunden bei 30C inkubiert. Die Reaktionsprodukte werden mit organischem Lösungsmittel wie Aceton oder Chloroform/Methanol (2:1) extrahiert und mittels HPLC bestimmt.

5

Besonders bevorzugt erfolgt die Bestimmung der Hydroxylase-Aktivität unter folgenden Bedingungen nach Bouvier, d'Harlingue und Camara (Molecular Analysis of carotenoid cyclae inhibition; Arch. Biochem. Biophys. 346(1) (1997) 53-64):

10

Der in-vitro Assay wird in einem Volumen von 250 □I Volumen durchgeführt. Der Ansatz enthält 50 mM Kaliumphosphat (pH 7.6),unterschiedliche Mengen an Organismenextrakt, 20 nM Lycopin, 250 □g an chromoplastidärem Stromaprotein aus Paprika, 0.2 mM NADP+, 0.2 mM NADPH und 1 mM ATP. NADP/NADPH und ATP werden in 10 ml Ethanol mit 1 mg Tween 80 unmittelbar vor der Zugabe zum Inkubationsmedium gelöst. Nach einer Reaktionszeit von 60 Minuten bei 30C wird die Reaktion durch Zugabe von Chloroform/Methanol (2:1) beendet. Die in Chloroform extrahierten Reaktionsprodukte werden mittels HPLC analysiert.

20

15

Ein alternativer Assay mit radioaktivem Substrat ist beschrieben in Fraser und Sandmann (Biochem. Biophys. Res. Comm. 185(1) (1992) 9-15).

25

Die Erhöhung der Hydroxylase-Aktivität kann durch verschiedene Wege erfolgen, beispielsweise durch Ausschalten von hemmenden Regulationsmechanismen auf Expressions- und Proteinebene oder durch Erhöhung der Genexpression von Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp.

30

Die Erhöhung der Genexpression der Nukleinsäuren kodierend eine Hydroxylase gegenüber dem Wildtyp kann ebenfalls durch verschiedene

Wege erfolgen, beispielsweise durch Induzierung des Hydroxylase-Gens durch Aktivatoren oder durch Einbringen von einer oder mehrerer Hydroxylase-Genkopien, also durch Einbringen mindestens einer Nukleinsäure kodierend eine Hydroxylase in denb Organismus der Gattung Blakesleaa.

5

In einer bevorzugten Ausführungsform erfolgt die Erhöhung der Genexpression einer Nukleinsäure kodierend eine Hydroxylase durch Einbringen von mindestens einer Nukleinsäure kodierend eine Hydroxylase in den Organismus der Gattung Blakesleaa.

10

Dazu kann prinzipiell jedes Hydroxylase-Gen, also jede Nukleinsäure, die eine Hydroxylase und jede Nukleinsäure, die eine β-Cyclase codiert, verwendet werden.

Bei genomischen Hydroxylase-Sequenzen aus eukaryontischen Quellen, 15

die Introns enthalten, sind für den Fall das der Wirtsorganismus nicht in der Lage ist oder nicht in die Lage versetzt werden kann, die entsprechende Hydroxylase zu exprimieren, bevorzugt bereits prozessierte Nukleinsäuresequenzen, wie die entsprechenden cDNAs zu verwenden.

20

25

Ein Beispiel für ein Hydroxylase-Gen ist eine Nukleinsäure, kodierend eine Hydroxylase aus Haematococcus pluvialis mit der Accession No. AX038729 (WO 0061764; Nukleinsäure: SEQ ID NO: 31, Protein: SEQ ID NO: 32), aus Erwinia uredovora 20D3 (ATCC 19321, Accession No. . D90087; Nukleinsäure: SEQ ID NO: 33, Protein: SEQ ID NO: 34) oder Hydroxylase aus Thermus thermophilus (DE 102 34 126.5) kodiert durch die Sequenz mit der SEQ ID NO 76.

30

Nummern: Hydroxylasen folgenden Accession sowie der CAA70427.1. CAA70888.1, CAB55625.1, Jemb|CAB55626.1, AF499108 1, AF315289 1, AF296158 1, AAC49443.1, NP_194300.1,

10

15

20

25

NP_200070.1, AAG10430.1, CAC06712.1, AAM88619.1, CAC95130.1, AAL80006.1, AF162276_1, AAO53295.1, AAN85601.1, CRTZ_ERWHE, CRTZ_PANAN, BAB79605.1, CRTZ_ALCSP, CRTZ_AGRAU, CAB56060.1, ZP_00094836.1, AAC44852.1, BAC77670.1, NP_745389.1, NP_344225.1, NP_849490.1, ZP_00087019.1, NP_503072.1, NP_852012.1, NP_115929.1, ZP_00013255.1

In den erfindungsgemäßen bevorzugten transgenen Organismen der Gattung Blakeslea liegt also in dieser bevorzugten Ausführungsform gegenüber dem Wildtyp mindestens ein weiteres Hydroxylase—Gen vor.

In dieser bevorzugten Ausführungsform weist die genetisch veränderte Organismus beispielsweise mindestens eine exogene Nukleinsäure, kodierend eine Hydroxylase oder mindestens zwei endogene Nukleinsäuren, kodierend eine Hydroxylase auf.

Bevorzugt verwendet man in vorstehend beschriebener bevorzugter Ausführungsform als Hydroxylase-Gene Nukleinsäuren, die Proteine kodieren, enthaltend die Aminosäuresequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76 oder eine von dieser Sequenz durch Substitution, Insertion oder Deletion von Aminosäuren abgeleitete Sequenz, die eine Identität von mindestens 30 %, vorzugsweise mindestens 50 %, bevorzugter mindestens 70%, noch bevorzugter mindestens 80 %, am bevorzugtesten mindestens 90%, insbesondere 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99% auf Aminosäureebene mit der Sequenz SEQ. ID. NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76 und die die enzymatische Eigenschaft einer Hydroxylase aufweisen.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich 30 beispielsweise aus verschiedenen Organismen, deren genomische Sequenz bekannt ist, wie vorstehend beschrieben, durch HomologievergleiWO 2004/063359 PCT/EP2004/000099

che der Aminosäuresequenzen oder der entsprechenden rückübersetzten Nukleinsäuresequenzen aus Datenbanken mit der SEQ ID. NO: 31, 33 oder 76 leicht auffinden.

Weitere Beispiele für Hydroxylasen und Hydroxylase-Gene lassen sich weiterhin beispielsweise ausgehend von der Sequenz SEQ ID NO: 31, 33 oder 76 aus verschiedenen Organismen deren genomische Sequenz nicht bekannt ist, wie vorstehend beschrieben, durch Hybridisierungs- und PCR-Techniken in an sich bekannter Weise leicht auffinden.

10

15

In einer weiter besonders bevorzugten Ausführungsform werden zur Erhöhung der Hydroxylase-Aktivität Nukleinsäuren in Organismen eingebracht, die Proteine kodieren, enthaltend die Aminosäuresequenz der Hydroxylase der Sequenz SEQ ID NO: 32, 34 oder kodiert durch die Sequenz mit der SEQ ID NO 76.

Geeignete Nukleinsäuresequenzen sind beispielsweise durch Rückübersetzung der Polypeptidsequenz gemäß dem genetischen Code erhältlich.

- 20 Bevorzugt werden dafür solche Codons verwendet, die entsprechend der Organismenspezifischen codon usage häufig verwendet werden. Die codon usage lässt sich anhand von Computerauswertungen anderer, bekannter Gene der betreffenden Organismen leicht ermitteln.
- In einer besonders bevorzugten Ausführungsform bringt man eine Nukleinsäure, enthaltend die Sequenz SEQ. ID. NO: 31, 33 oder 76 in den Organismus ein.
- Alle vorstehend erwähnten Hydroxylase-Gene sind weiterhin in an sich bekannter Weise durch chemische Synthese aus den Nukleotidbausteinen wie beispielsweise durch Fragmentkondensation einzelner überlappender,

komplementärer Nukleinsäurebausteine der Doppelhelix herstellbar. Die chemische Synthese von Oligonukleotiden kann beispielsweise, in bekannter Weise, nach der Phosphoamiditmethode (Voet, Voet, 2. Auflage, Wiley Press New York, Seite 896-897) erfolgen. Die Anlagerung synthetischer Oligonukleotide und Auffüllen von Lücken mithilfe des Klenow-Fragmentes der DNA-Polymerase und Ligationsreaktionen sowie allgemeine Klonierungsverfahren werden in Sambrook et al. (1989), Molecular cloning: A laboratory manual, Cold Spring Harbor Laboratory Press, beschrieben.

10

5

Der in der Transformation (i) eingesetzte Vector umfasst daher in weiteren Ausführungsformen der Erfindung bevorzugterweise eine Sequenz codierend für eine Hydroxlase, insbesondere eine Hydroxlase aus Haematococcus pluvialis mit der SEQ ID NO: 70 oder eine Hydroxlase aus Erwinia uredova mit der SEQ ID NO: 71. oder eine Hydroxylase aus Thermus thermophilus kodiert durch die Sequenz mit der SEQ ID NO 76.

Vorzugsweise wird durch die Transformation das Gen der Phytoendesaturase ausgeschaltet.

20

15

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise ferner die Expression regelnde und unterstützende Bereiche, insbesondere Promotoren und Terminatoren.

25

30

Der in der Transformation (i) eingesetzte Vector enthält vorzugsweise den gpd und/oder den ptef1 Promotor und/oder den trpC Terminator. Diese haben sich zur Transformation der Blakeslea besonders bewährt. Auch der Einsatz von dem Fachmann geläufigen "inverted repeats" (IR, Römpp Lexikon der Biotechnologie 1992, Thieme Verlag Stuttgart, Seite 407 "Invers repetitive Sequenzen") zur Regelung der Expression bzw. Transkription liegt im Rahmen der Erfindung.

15

20

Vorteilhafterweise weist der im Vector eingesetzte gpd Promotor die Sequenz SEQ ID NO: 1 auf. Vorteilhafterweise weist der im Vector eingesetzte trpC Terminator die Sequenz SEQ ID NO: 2 auf. Vorteilhafterweise weist der im Vector eingesetzte ptef1 Promotor die Sequenz SEQ ID NO: 35 auf.

Insbesondere werden dabei der gpd Promotor und der trpC Terminator aus Aspergillus nidulans und der ptef1 Promotor aus Blakeslea trispora eingesetzt.

Insbesondere enthält der in der Transformation (i) eingesetzte Vector ein Resistenzgen. Bevorzugterweise handelt es sich um ein Hygromycin-Resistenzgen (hph), insbesondere eines aus E. coli. Dieses Resistenzgen hat sich bei dem Nachweis der Transformation und Selektion der Zellen als besonders geeignet herausgestellt.

Als Promotor für hph wird also bevorzugt p-gpdA, der Promotor der Glycerinaldehyd-3-phosphatdehydrogenase aus Aspergillus nidulans genutzt. Als Terminator für hph wird bevorzugt t-trpC, der Terminator des Gens trpC, codierend für Anthranilatsynthasekomponenten aus Aspergillus nidulans genutzt.

Als Vectoren haben sich Abkömmlinge des pBinAHyg Vectors als besonders geeignet herausgestellt. Der zur Transformation eingesetzte Vector umfasst also bevorzugterweise die SEQ ID NO: 3.

Hinzu kommen je nach gewünschtem Carotinoid oder dessen Vorstufe eine Sequenz codierend für eine Hydroxylase, Ketolase, Phytoendesaturase usw. wie diese zuvor beschrieben wurden. Die Vectoren umfassen also in einer Ausführugsform der Erfndung die Sequenz SEQ ID NO: 69 codierend für die Phytoendesaturase. Die Vectoren umfassen ferner in

einer weiteren Ausführugsform der Erfndung die Sequenz SEQ ID NO: 72 codierend für eine Ketolase. Die Vectoren umfassen weiter in einer weiteren Ausführugsform der Erfndung die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydoxylase. Entsprechende Kombinationen der zuvorgenannten Sequenzen liegen ebenso im Rahmen der Erfindung. So umfasst der Vector in einer Ausführungsform sowohl eine Sequenz SEQ ID NO: 72 codierend für eine Ketolase als auch die Sequenz SEQ ID NO: 70 oder 71 oder 76 codierend für eine Hydoxylase und ermöglicht so die Herstellung von Astaxanthin.

10

15

20

30

Insbesondere sind Vectoren ausgewählt aus der Gruppe bestehend aus den SEQ ID NO: 37 bis 51 und 62 im Rahmen der Erfindung einsetzbar.

Die genetisch veränderten Organismen können zur Produktion von Carotinoiden, Xanthophyllen oder deren Vorstufen, insbesondere Bixin, Phytoen, Astaxanthin, Zeaxanthin und Canthaxanthin verwendet werden. Auch können neue, im Wildtyp natürlicherweise nicht vorkommende Carotinoide durch Einbringung der entsprechenden genetischen Information von den gezielt genetisch veränderten Zellen bzw. dem durch sie gebildeten Mycel erzeugt und anschließend isoliert werden.

Die gentechnisch verändertern Zellen werden nach der Selektion kultiviert, so daß Carotinoiden oder deren Vorstufen bereitgestellt werden können.

Bevorzugterweise ist die Gewinnung von Carotinoiden oder deren Vorstufen mit den gezielt genetisch veränderten Zellen bzw. das durch sie gebildete Mycel möglich.

Die Kultivierung der Organismen unterliegt keinen Besonderheiten. Vorteilhafterweise werden, insbesondere bei der Verwendung von Blakeslea

WO 2004/063359

trispora, entgegengesetzte Paarungstypen gemeinsam kultiviert, da dies zu besserem Wachstum und Produktion führt.

Wird die gentechnische Veränderung nur in Zellen eines der vorkommenden Paarungstypen (bei Blakeslea trispora (+) oder (-)) durchgeführt, so wird zur Kultivierung der entsprechend andere, nicht veränderte Paarungstyp zugesetzt, da so eine gute Produktion der Carotinoide oder deren Vorstufen aufgrund der von dem zweiten, nicht veränderten Paarungstyp abgegebenen Substanzen (z. B. Trisporsäuren) zu erreichen ist. Vorteilhafterweise wird jedoch die gentechnische Veränderung in Zellen beider Paarungstypen vorgenommen und diese zusammen kultiviert. Hierdurch wird ein besonders gutes Wachstum und eine optimale Produktion der Carotinoiden oder deren Vorstufen erreicht. Auch eine (künstliche) Zugabe der Trisporsäuren ist möglich und sinnvoll.

15

20

5

Trisporsäuren sind Sexualhormone in Mucorales Pilzen, wie Blakeslea, welche die Bildung von Zygophoren und die Produktion von β -Carotin stimulieren (van den Ende 1968, J. Bacteriol. 96:1298 - 1303, Austin et al. 1969, Nature 223:1178 – 1179, Reschke Tetrahedron Lett. 29:3435 – 3439, van den Ende 1970, J. Bacteriol. 101:423 – 428).

25

30

Es können alle dem Fachmann geläufigen Medien eingesetzt werden, so weit sich diese zur Kultivierung der eingesetzten Organismen und deren Carotinoid Produktion eigenen. Insbesondere müssen bei Einsatz der GVO keine Carotinoidbiosynthese Inhibitoren eingesetzt werden. Die eingesetzten Medien beinhalten vorzugsweise Zusätze, wie eine oder mehrere Kohlenstoffquellen, eine oder mehrere Stickstoffquellen, Mineralsalze und Thiamine. Bevorzugterweise werden Zusätze eingesetzt, wie sie aus der WO 03/038064 A2, Seite 4, Zeile 30 bis Seite 5 Zeile 7 hervorgehen. Besonders bevorzugt wird als Kohlenstoffquelle Glukose und als Stick-

10

15

20

25

30

stoffquelle Asparagin, pflanzliche oder tierische Extrakte, wie Baumwollsatöl, Sojaöl, Baumwollsamenmehl oder Hefe-Extrakt zugesetzt.

Die Kultivierung kann entweder unter aeroben oder anaeroben Bedingungen durchgeführt werden. Auch eine gemischte, zunächst aerobe und anschließend anaerobe Kultivierung, wie sie aus der DE 101 30 323 bekannt ist, ist möglich. Temperatur und Luftfeuchtigkeit werden dabei jeweils zum optimalen Wachstum eingestellt. Bevorzugterweise liegt die Temperatur bei der Kultivierung zwischen ca. 20 und ca. 34 °C, insbesondere zwischen ca. 26°C und ca. 28°C. Die Kultivierung kann ferner kontinuierlich, batch- oder satzweise erfolgen.

Die Kultivierung erfolgt vorzugsweise bis zu einem Feststoffgehalt zwischen etwa 1 und etwa 20 %, bevorzugt 3 und 15 % und besonders bevorzugt 4 und 11 %. Insbesondere ist wichtig dass die Kulturbrühe pumpbar bleibt, so dass sie in den nachfolgenden Verfahrensschritten ver- und bearbeitbar bleibt. Ist der Feststoffgehalt zu klein, so muss ein großer Aufwand bei der Aufkonzentrierung oder Trocknung betrieben werden.

Die Kultivierung bzw. Fermentation kann in den üblichen Apparaturen durchgeführt werden. Hierzu kommen alle für die jeweils eingesetzten Mikroorganismen und deren Produkte geeigneten Apparaturen in Betracht. Insbesondere solche, wie sie aus dem Römpp Lexikon Biotechnologie (1992 Georg Thieme Verlag, Stuttgart) unter dem Stichwort "Bioreaktor" auf Seiten 123 - 126 angegeben sind. Besonders bevorzugt ist der Einsatz von Rührkesselreaktoren mit versch. Einbauten, Blasensäulen verschiedener Bauarten, etc.

Die nach dem erfindungsgemäßen Verfahren bereitgestellten Carotinoiden oder deren Vorstufen, insbesondere Bixin, Phytoen oder Xanthophylle, besonders bevorzugt Astaxanthin oder Zeaxanthin eignen sich besonders

zur Herstellung von Zusätzen für Futter-, Nahrungs- und Nahrungsergänzungsmittel, kosmetischen, pharmazeutischen oder dermatologischen Zubereitungen.

Die Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe aus der Kultur der gentechnisch veränderten Mikroorganismen erfolgt nach zwei Varianten a) oder b), becorzugt ist auch eine Kombination aus a) und b);

10

15

a:

- 1) Abtrennung der Biomasse,
 - IA) ggf. Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser,
 - IB) Sterilisation und Zellaufschluß der Biomasse,
 - IC) ggf. Trocknung und/oder homogene Verteilung und
- partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse,
 IIA)

 Entfernung von Lösemittelresten aus der Carotinoidhaltigen Biomasse,

- ggf. homogene Suspension der Biomasse mit einem
 Biomasse-Feststoffgehalt > 2 % und < 50 %, und
- Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels,

IIB)

 Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration;

20

25

oder b): Homogene Suspendierung der Feststoffe der Kulturbrühe I) und 5 bei einem Feststoffgehalt der Kulturbrühe von > 2 % IIA) ggf. Konzentration der Kulturbrühe auf einen Fest-1) stoffgehalt < 50 % und Trocknung der Kulturbrühe zur Herstellung des Nah-2) 10 rungsmittels oder bei einem Feststoffgehalt von < 2 % der Kulturbrühe, IIB) Konzentration der Kulturbrühe auf einen Feststoffge-1) 15 halt > 2 % und < 50 % und 2) Trocknung der Suspension zur Herstellung des Nahrungsmittels, oder 20 unabhängig vom Feststoffgehalt der Kulturbrühe, IIC) Abtrennung der Biomasse, 1) ggf. Waschen der Biomasse mit Carotinoide nicht lö-2) senden Lösungsmitteln, insbesondere Wasser, Sterilisation und Zellaufschluß, 3) 25 agf. Trocknung und homogene Verteilung, 4) partielle Extraktion der Carotinoide aus der Biomasse 5) mittels eines Carotinoide lösendes Lösungsmittels, Abtrennung der Carotinoid-haltigen Biomasse vom 5a) Carotinoid-haltigen Lösungsmittel,

- 5b) Entfernung von Lösemittelresten aus der Biomasse und
- 5c) Trocknung der Biomasse zur Herstellung des Nahrungsmittels,
- 6) Kristallisation der Carotinoide aus dem in 5a) verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- Die erfindungsgemäße Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe aus der Kultur der gentechnisch veränderten Mikroorganismen erfolgt nach zwei Varianten a) oder b) ermöglicht die gleichzeitige Herstellung von zwei Produkten.

20

25

5

- Durch die erfindungsgemäße Kombination der Herstellung von zwei Produkten, insbesondere bei der Bereitstellung gemäß Variante a), nämlich dem mindestens einem Carotinoid und dem Carotinoid-haltigem Nahrungsmittel ist keine vollständige Extraktion der Carotinoide aus der Biomasse nötig, so dass der Aufwand bei der Extraktion geringer ausfällt. Das Carotinoid muss trotz vollständiger Verwertung nur partiell extrahiert werden, ohne dass es zu Produktverlusten kommt. Dies bedingt geringere Lösungsmittelmengen und damit einhergehend einen geringeren Aufwand bei den Maßnahmen zu deren Wiederverwendung. Zudem werden Abfälle weitestgehend vermieden, da die Biomasse nicht als Abfall anfällt, sondern zum hochwertigen Nahrungsmittel weiterverarbeitet wird. Somit ergeben sich geringere Kosten für die Verfahren durch Ausnutzen von Synergien.
- Nach dem erfindungsgemäßen Verfahren mit Bereitstellung gemäß Variante b) erhältliche Nahrungsmittel enthalten also bereits nach der Herstel-

10

15

20

25

30

lung große Mengen an Carotinoiden, die nicht zugesetzt werden müssen. Dadurch, dass das Nahrungsmittel neben dem mindestens einen Carotinoid auch Blakeslea trispora enthält ist sein Nährstoffgehalt zudem gesteigert. Insbesondere ist der Nährstoffgehalt nach den bevorzugten Alternativen IIA und IIB stark gesteigert, da es neben dem mindestens einen Carotinoid und Blakeslea trispora zusätzlich alle Medienbestandteile der Fermentation enthält. Ferner benötigt das Verfahren keine zusätzlichen, aufwendigen Aufarbeitungs- und Herstellungsschritte, sondern die homogenisierte und ggf. entwässerte Blakeslea trispora-haltige Kulturbrühe kann direkt ohne Umwege zur Herstellung des Nahrungsmittels getrocknet werden. Es fallen demnach praktisch keine Abfälle an, abgesehen vom wässrigem Medium bei der Alternative IIB, welches jedoch unproblematisch in einer Kläranlage gereinigt werden kann. Zusätzlich wird in allen drei Alternativen die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da gemäß IIA und IIB keine verlustreichen Trenn- bzw. Aufarbeitungsschritte vorgenommen werden müssen. Bei der Alternative IIIC wird ebenfalls die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da ein Teil in der Biomasse zum Nahrungsmittel verarbeitet wird und der andere Teil zur Gewinnung reiner Carotinoide extrahiert wird. Durch die erfindungsgemäße Kombination der Herstellung von zwei Produkten gemäß IIC, nämlich dem Carotinoid-haltigem Nahrungsmittel und den Carotinoiden an sich, ist vorteilhaft, dass wiederum im wesentlichen keine Abfälle entstehen und eine vollständige Extraktion der Carotinoide aus der Biomasse unnötig ist, so dass sonst bei der Extraktion anfallende Aufwand geringer ausfällt. Das oder die wertvollen Carotinoide müssen trotz vollständiger Verwertung nur partiell extrahiert werden, ohne dass es zu Produktverlusten kommt. Dies bedingt geringere Lösungsmittelmengen und damit einhergehend einen geringeren Aufwand bei den Maßnahmen zu deren Wiederverwendung. Zudem werden Abfälle weitestgehend vermieden, da die Biomasse nicht als Abfall anfällt, sondern zum hochwertigen Nahrungsmittel verarbeitet

10

15

20

25

30

wird. Somit ergeben sich geringere Kosten für die Verfahren durch Ausnutzen von Synergien.

Unter "hochrein" soll in der vorliegenden Anmeldung eine Reinheit des mindestens einen Carotinoids von mindestens 95%, bevorzugt > 95%, vorzugsweise > 96%, besonders bevorzugt > 97%, ganz besonders bevorzugt > 98%, höchst bevorzugt > 99% verstanden werden.

Als nach dem erfindungsgemäßen Verfahren herstellbare Carotinoide kommen alle natürlichen und künstlichen Carotine und Xanthophylle in Betracht. Insbesondere ist das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β-Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β-Carotin, Lutein, Phytofluen, Bixin und Phytoen ausgewählt. Bevorzugterweise handelt es sich um Astaxanthin oder Zeaxanthin. Die Carotinoide können nach dem erfindungsgemäßen Verfahren einzeln oder als Gemische von zwei oder mehrerer der zuvor genannten Carotinoide erhalten werden. Insbesondere bei Einsatz der weiter unten angegebenen gentechnisch veränderten Organismen (GVO) kann das oder können die Carotinoide gezielt hergestellt werden.

Als Nahrungsmittel werden Zusammensetzungen angesehen, die der Ernährung dienen. Darunter fallen auch Zusammensetzungen für die Ergänzung der Ernährung. Insbesondere werden als Nahrungsmittel Tierfuttermittel und Tierfutterergänzungsmittel angesehen.

Nach der Kultivierung kann, gemäß Variante a) der Bereitstellung, die Biomasse von der Kulturbrühe abgetrennt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden. Hierunter fallen insbesondere die mechanischen Verfahren, wie Filtration und Zentrifugation, die auf der

15

20

25

30

Ausnutzung von Schwerkraft, Zentrifugalkraft, Druck oder Vakuum beruhen. Zu den einsetzbaren Verfahren und Apparaten gehören daneben u. a. Querstromfiltration bzw. Mebrantechniken wie Osmose, umgekehrte Osmose, Mikrofiltration, Ultrafiltration, Nanofiltration, Kuchenfiltrationsverfahren (z.B. mittels Pressfilterautomaten, (Membran-, Rahmen- oder Kammer-)Filterpressen, (Rühr-)drucknutschen, Saugnutschen, (Vakuum-)bandfiltern, (Vakuum-)Trommelfiltern, Drehfiltern, Kerzenfiltern), Zentrifugationsverfahren mittels kontinuierlich oder diskontinuierlich betriebener Zentrifugen oder Filterzentrifugen (z.B. Stülpfilterzentrifugen, Schälzentrifugen, Schubzentrifugen, Siebschneckenzentrifugen, Gleitzentrifugen, Separatoren oder Dekantern), Verfahren unter Ausnutzung der Schwerkraft wie Flotation, Sedimentation, Sink-Schwimm-Aufbereitung und Klären. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekanters oder durch Filtration, mittels einer Mebranfiltrationseinheit durchgeführt.

In dem zweiten Schritt der Bereitstellung nach Variante b) wird eine homogen verteilte Suspension der Feststoffe in der Kulturbrühe erzeugt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden verwendet werden. Insbesondere kommen dazu Dispergiergeräte, wie ein Ultra-Turrax® (im Labormaßstab) zum Einsatz. Ein Zellaufschluss ist nicht notwendig, kann aber vorgenommen werden.

Falls notwendig, kann die Kulturbrühe entwässert werden, um einen geeigneten Feststoffgehalt zwischen > 2 % und < 50 % zu erreichen. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden. Hierunter fallen insbesondere die mechanischen Verfahren, wie Filtration und Zentrifugation, die auf der Ausnutzung von Schwerkraft, Zentrifugalkraft, Druck oder Vakuum beruhen. Zu den einsetzbaren Verfahren und Apparaten gehören daneben u. a. Querstromfiltration bzw. Mebrantechniken wie Osmose,

umgekehrte Osmose, Mikrofiltration, Ultrafiltration, Nanofiltration, Kuchenfiltrationsverfahren (z.B. mittels Pressfilterautomaten, (Membran-, Rahmen- oder Kammer-)Filterpressen, (Rühr-)drucknutschen, Saugnutschen, (Vakuum-)bandfiltern, (Vakuum-)Trommelfiltern, Drehfiltern, Kerzenfiltern), Zentrifugationsverfahren mittels kontinuierlich oder diskontinuierlich betriebener Zentrifugen oder Filterzentrifugen (z.B. Stülpfilterzentrifugen, Schälzentrifugen, Schubzentrifugen, Siebschneckenzentrifugen, Gleitzentrifugen, Separatoren oder Dekantern), Verfahren unter Ausnutzung der Schwerkraft wie Flotation, Sedimentation, Sink-Schwimm-Aufbereitung und Klären. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekanters oder wird durch Filtration, mittels einer Mebranfiltrationseinheit durchgeführt. Anschließend wird die Kulturbrühe getrocknet. Hierzu können wiederum alle dem Fachmann bekannten Verfahren und Apparate eingesetzt werden Insbesondere eignen sich Apparate zur thermischen Trocknung wie Konvektions-, Kontakt- und Strahlungstrocknung, z.B. Horden-, Kammer-, Kanal-, Flachbahn-, Teller-, Drehtrommel-, Rieselschacht-, Siebband-, Strom-, Wirbelschicht-, Fließbett-, Schaufel-, Kugelbett-, , Heizteller-, Dünnschicht-, Walzen-, Band-, Siebtrommel-, Schnecken-, Taumel-, Kontakt-Scheiben-, Infrarot-, Mikrowellen- und Gefriertrockner, Sprühtrockner oder Sprühtrockner mit integrierter Wirbelschicht, die durch Dampf, Öl, Gas oder elektrischen Strom ggf. beheizt und ggf. unter Vakuum betrieben werden. Die Betriebsweie kann dabei je nach Apparat kontinuierlich oder diskontinuierlich sein. Daneben oder damit in Kombination können die oben bereits angegeben mechanischen Verfahren zur Fest/flüssig-Trennung verwendet werden.

Eine Granulierung durch Extrusion wie dies aus der WO 97/36996 Å2 hervorgeht ist jedoch nicht notwendig. Durch die Trocknung wird das Nahrungsmittel haltbar und lagerfähig.

10

15

20

25

Insbesondere wird die Kulturbrühe sprühgetrocknet. Bevorzugt wird zur Trocknung die Sprühtrocknung eingesetzt, wie sie aus der DE 101 04 494 A1, DE-A-12 11 911 oder EP 0 410 236 A1 bekannt sind. Ergänzend wird auf vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Sprühtrocknung" und Römpp Lexikon Biotechnologie, Georg Thieme Verlag, 1992, "Zerstäubungstrocknung" verwiesen. Die Sprühtrocknung bietet den Vorteil der kurzen Verweilzeit des Produkts in der heißen Zone des Trockners, so dass eine besonders schonende Trocknung erzielt wird.

10

Bei der Sprühtrocknung werden Eingangstemperaturen von ca. 115°C – 180°C, bevorzugt 120°C –130°C, und Ausgangstemperaturen von ca. 50°C – 80°C, bevorzugt 55°C – 70°C gewählt. Als Trocknungsgas wird vorzugsweise Stickstoff eingesetzt.

15

20

Gegebenenfalls können zur Erzielung einer besseren Rieselfähigkeit Rieselhilfsmittel, wie Kieselsäuren etc. zugesetzt werden. Der Einsatz von inerten Trägermaterialien, d.h. niedermolekularen anorganischen Trägern wie NaCl, CaCO3, Na2SO4 oder MgSO4, organischen Trägern wie Glucose, Fructose, Saccharose, Dextrine oder Stärkeprodukten (Roggen-, Gersten-, Hafermehl, Weizengrießkleie) ist denkbar.

Das getrocknete Produkt weist bevorzugt eine Restfeuchte von weniger als 10 %, bevorzugt weniger als 5 % bezogen auf die Trockenmasse auf. Sein Carotinoidgehalt liegt zwischen 0,05 und 20 %, insbesondere 1 und 10 % bezogen auf die Trockenmasse.

25

Das so hergestellte Nahrungsmittel kann entweder direkt verwendet werden oder mittels weiterer Zusätze aufbereitet werden, so wie dies ebenfalls aus der DE 101 04 494 A1 bekannt ist.

WO 2004/063359 PCT/EP2004/000099

51

Gemäß der Alternative IIC wird nach der Kultivierung und vor der Trocknung der Biomasse, die Biomasse von der Kulturbrühe zunächst abgetrennt. Hierzu können alle dem Fachmann geläufigen und üblicherweise einsetzbaren Methoden zur fest/flüssig-Trennung eingesetzt werden, wie sie bereits oben bei der Entwässerung genannt wurden. Bevorzugterweise erfolgt die Abtrennung der Biomasse von der Kulturbrühe durch Zentrifugation mittels eines Dekanters oder durch Membranfiltration durchgeführt.

5

20

25

30

Anschließend erfolgt optional das Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser, wodurch insbesondere Wasser lösliche Komponenten entfernt werden. Dieser Schritt
kann gegebenenfalls unter Verwendung weiterer Carotinoide nicht lösenden Lösungsmittel (z. B. Alkohole) ergänzt werden, was aber im Rahmen
der Erfindung nicht notwendig ist und zur Vermeidung von Abfällen nicht
bevorzugt ist.

Anschließend erfolgen die Sterilisation und der sich anschließende oder gleichzeitige Zellaufschluß der Zellen in der Biomasse. Durch die Sterilisation werden die Mikroorganismen abgetötet und gegebenenfalls vorhandene Enzymaktivität beendet. Dies ist zur Verhinderung des Abbaus der Biomasse bzw. der darin enthaltenen Stoffe, insbesondere der Carotinoide und für die Haltbarkeit von Bedeutung.

Die Sterilisation kann mit einem üblichen, dem Fachmann geläufigen Verfahren durchgeführt werden. Hierzu gehören die Sterilisation mittels Dampf, insbesondere bei Temperaturen größer 120 °C unter Druck (≥ 1 bar) und Zeitdauern von ≥ ca. 20 min. sowie die Behandlung mit energiereichen Strahlen, wie UV-, Mikrowellen, Gamma- oder Beta-Strahlen. Bevorzugterweise erfolgt die Sterilisation im Rahmen des erfindungsgemäßen Verfahrens mittels Dampf oder Mikrowellenstrahlung.

WO 2004/063359 PCT/EP2004/000099

Durch den nachfolgenden oder gleichzeitigen Zellaufschluß, werden die innerhalb der Zellen vorliegenden Carotinoide freigesetzt. Der Zellaufschluß kann ebenfalls mit allen dem Fachmann bekannten üblichen Verfahren erfolgen. Hierzu gehören mechanische und nicht mechanische Methoden. Zu den mechanischen Methoden zählen Trockenmahlen, Naßmahlen, Rühren, Homogenisieren (z.B. im Hochdruckhomogenisator) und die Verwendung von Ultraschall oder Mikrowellen. Als nicht mechanische Methoden kommen physikalische, chemische und biochemische Methoden in Betracht. Hierzu gehören Kurzzeiterhitzen, Kurzzeitgefrieren, Osmotischer Schock, Trocknung, Behandlung mit Säuren oder Laugen sowie ein enzymatischer Aufschluss. Günstiger Weise wird zum Zellaufschluss jedoch das zur Sterilisation verwendete Verfahren eingesetzt. Bevorzugterweise wird also ebenfalls der Zellaufschluss mittels Dampf oder Mikrowellen Strahlung durchgeführt.

15

20

10

Die Sterilisation und/oder der Zellaufschluss können kontinuierlich oder diskontinuierlich durchgeführt werden.

Die Sterilisation und/oder der Zellaufschluss können im zur Kultivierung eingesetzten Bioreaktor oder in anderen Apparaturen, wie Autoklaven usw. durchgeführt werden. Bei kontinuierlicher Durchführung kann das aus der WO 01/83437 A1 bekannte Mikrowellen verwendende Verfahren und entsprechende Apparaturen eingesetzt werden.

Vor der Extraktion wird die Biomasse gegebenenfalls getrocknet und/oder homogenisiert. Hierzu k\u00f6nnen wiederum alle dem Fachmann bekannten \u00fcblichen Verfahren und Ger\u00e4te eingesetzt werden. Insbesondere eignen sich Apparate zur thermischen Trocknung wie Konvektions-, Kontakt- und Strahlungstrocknung, z.B. Horden-, Kammer-, Kanal-, Flachbahn-, Teller-, Drehtrommel-, Rieselschacht-, Siebband-, Strom-, , Wirbelschicht-, Fließbett-, Schaufel-, Kugelbett-, , Heizteller-, D\u00fcnnschicht-, Walzen-, Band-,

15

20

Siebtrommel-, Schnecken-, Taumel-, Kontakt-Scheiben-, Infrarot-, Mikrowellen- und Gefriertrockner, Sprühtrockner oder Sprühtrockner mit integrierter Wirbelschicht, die durch Dampf, Öl, Gas oder elektrischen Strom ggf. beheizt und ggf. unter Vakuum betrieben werden. Die Betriebsweise kann dabei je nach Apparat kontinuierlich oder diskontinuierlich sein. Daneben oder damit in Kombination können die oben bereits angegeben mechanischen Verfahren zur Fest/flüssig-Trennung verwendet werden.

Eine Granulierung durch Extrusion wie dies aus der WO 97/36996 A2 hervorgeht ist jedoch nicht notwendig.

Anschließend erfolgt die partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse. Sowohl in dem Lösungsmittel als auch in der Biomasse sind nun Carotinoide enthalten, wobei sich in dem Lösungsmittel bevorzugterweise der Großteil der Carotinoide befindet.

Aus dem Lösungsmittel werden anschließend die hochreinen Carotinoide isoliert, wohingegen die Biomasse zu einem hochwertigen, Carotinoidhaltigen Nahrungsmittel weiterverarbeitet wird, welches durch den vorhergehenden Zellaufschluss auch eine gute Bioverfügbarkeit der Carotinoide aufweist.

- Unter partieller Extraktion soll demnach die bewusst unvollständige Extraktion der Carotinoiden aus der Biomasse verstanden werden (vgl. oben). Bevorzugterweise wird im Rahmen der Erfindung durch die Extraktion also weniger als 100 % der in der Biomasse enthaltenen Gesamtmenge der Carotinoide aus dieser extrahiert.
- Dies ist von großem Vorteil, da der Aufwand zur Extraktion mit der abnehmenden Menge Carotinoid in der Biomasse überproportional zunimmt.

10

15

20

Zur Extraktion werden Lösungsmittel eingesetzt, die Carotinoide lösen, wie z. B. Hexan, Ethylacetat, Dichlormethan oder überkritisches Kohlendioxid. Bevorzugterweise wird erfindungsgemäß als Lösungsmittel Dichlormethan oder überkritisches Kohlendioxid eingesetzt, wobei beim Einsatz von überkritischem Kohlendioxid die darin enthaltenen Carotinoide anschließend in Dichlormethan überführt werden können oder das Wertprodukt direkt durch Entspannung des Kohlendioxids gewonnen werden kann. Dabei werden die Mengen der Lösungsmittel und Durchmischungszeiten derart gewählt, dass die gewünschte Menge Carotinoide aus der Biomasse extrahiert wird. Insbesondere wird der Extraktionsschritt nur einmal durchgeführt, was technisch und wirtschaftlich sinnvoll ist (vgl. oben).

Zur Durchführung der Extraktion können alle üblichen Verfahren und Apparaturen eingesetzt werden. Insbesondere wird bei nicht getrockneter, aber aufgeschlossener Biomasse eine flüssig/flüssig (Carotinoid liegt in flüssigen Zellbestandteilen gelöst vor und wird daraus extrahiert und bei getrockneter Biomasse eine fest/flüssig Extraktion durchgeführt. Es können Kalt- und Heißextraktion in bestimmten Temperaturbereichen, sowohl kontinuierliche (z.B. Soxhlet-Extraktion, Perforation und Perkolation) als auch diskontinuierliche Verfahren, zu denen beispielsweise Ausschütteln, Auslaugen, Auskochen und Digerieren gehören, verwendet werden. Sie können auch im Gegenstromverfahren durchgeführt werden.

Für die flüssig/flüssig Extraktion können beispielsweise Blasensäulen, , pulsierende Kolonnen, Kolonnen mit rotierenden Einbauten, Mixer-Settler-Batterien oder Rührkessel usw. verwendet werden.

Die fest/flüssig Extraktion kann mittels üblicher Apparaturen durchgeführt werden. Vorzugsweise werden Rührkessel oder Mixer-Settler-Apparate eingesetzt.

Alternativ kann der Zellaufschluß ohne vorherige Abtrennung des Fermentaionsmediums erfolgen und sich dann eine direkte Trennung einer sich bildenden Carotinoidsuspension von der Biomasse z. B. mittels eines Dekanters durchgeführt werden. Anschließend wird die Carotinoidsuspension in Dichlormethan aufgenommen und weiterverarbeitet oder alternativ durch Wäschen mit verschiedenen wässrigen Lösungen aufgereinigt.

Zur Isolierung der hochreinen Carotinoide aus dem Lösungsmittel wird eine Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration durchgeführt. Die verbleibende Mutterlauge kann nach Destillation dem Verfahren erneut zugeführt werden, so dass Produktverluste trotz geringem Aufwand minimiert werden.

15

30

10

5

Die Kristallisation kann wie üblich erfolgen. Ergänzend wird auf vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Kristallisation" verwiesen.

Bevorzugterweise erfolgt die Kristallisation durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel. Es wird also kontinuierlich die Löslichkeit der Carotinoide erniedrigt, bis diese als reine Kristalle ausfallen. Hierbei wird vorzugsweise ein "niederer Alkohol" oder Wasser verwendet. Als niederer Alkohol werden aliphatische Alkohole mit 1 bis 4 Kohlenstoffatomen angesehen. Hierzu gehören Methanol, Ethanol, Propanol, Isopropanol, 1-Butanol, tert.-Butanol und sec.-Butanol. Bevorzugterweise wird Methanol eingesetzt.

Die Carotinoid-Lösung kann dabei erwärmt werden, wobei die Temperatur vorzugsweise < 100 °C, insbesondere < 60 °C gehalten wird, so dass Dichlormethan abdestilliert wird. Auch der Einsatz von Vakuum ist denk-

bar. Anschließend werden die CarotinoidKristalle isoliert, welches durch übliche Maßnahmen, insbesondere durch Filtration erfolgen kann. Es können sich, falls gewünscht, weitere optionale Trocknungs- und/oder Reinigungsschritte anschließen. Notwendig sind diese jedoch nicht, da die Carotinoid Kristalle bereits hochrein sind.

Die Carotinoide fallen als hochreine Kristalle an und weisen eine Reinheit von mindestens 95%, bevorzugt > 95%, vorzugsweise > 96%, besonders bevorzugt > 97%, ganz besonders bevorzugt > 98%, höchst bevorzugt > 99% auf.

Die erzielbaren Ausbeuten liegen zwischen 45% und 95%, bevorzugt zwischen 70% und 95% bezogen auf die in der Kulturbrühe vorliegende Menge (0.5 - 15 g/L), bevorzugt 1 - 10 g/L).

15

20

25

30

10

5

Zur Weiterverarbeitung der ebenfalls Carotinoid-haltigen Biomasse zu einem hochwertigen Nahrungsmittel wird zunächst eine Entfernung von Lösemittelresten aus der Carotinoid-haltigen Biomasse vorgenommen. Hierzu erfolgt bevorzugterweise eine Wasserdampfdestillationen bzw. ein so genanntes Strippen mit Wasserdampf (vgl. Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Strippen").

Danach kann gegebenenfalls die Biomasse in der oben abgetrennten Kulturbrühe homogen suspendiert werden, wobei ein Feststoffgehalt > 100 g/L und < 600 g/L eingehalten werden sollte, so daß die nachfolgende Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels ohne technische Schwierigkeiten erfolgen kann. D.h. die Suspension muß pumpbar sein. Als Trocknungsverfahren kommen alle bereits genannten Verfahren und Apparaturen in Frage. Insbesondere wird zur Trocknung die Sprühtrocknung eingesetzt. Dabei kann wie aus der DE 101 04 494 A1 bekannt verfahren werden.

Bei der Sprühtrocknung werden Eingangstemperaturen von ca. 100°C – 180°C, bevorzugt 120°C –130°C, und Ausgangstemperaturen von ca. 50 – 80°C, bevorzugt 55°C – 70°C gewählt. Als Trocknungsgas wird vorzugsweise Stickstoff eingesetzt.

Das so hergestellte Nahrungsmittel kann entweder direkt verwendet werden oder mittels weiterer Zusätze aufbereitet werden, so wie dies ebenfalls aus der DE 101 04 494 A1 bekannt ist.

10

15

20

25

30

Als Nahrungsmittel werden Zusammensetzungen angesehen, die der Ernährung dienen. Darunter fallen auch Zusammensetzungen für die Ergänzung der Ernährung. Insbesondere werden als Nahrungsmittel Tierfuttermittel und Tierfutterergänzungsmittel angesehen. Ergänzend wird auf Römpp Lexikon Chemie CD-ROM Version 2.0, Georg Thieme Verlag, 1999, "Nahrungsmittel" verwiesen.

Das trockene Produkt weist bevorzugt eine Restfeuchte von weniger als 5 % bezogen auf die Trockenmasse auf. Sein Carotinoidgehalt liegt zwischen 0,05 und 20 %, insbesondere 1 und 10 % bezogen auf die Trokkenmasse. Der gewünschte Carotinoidgehalt ist über das Ausmaß der Extraktion steuerbar (vgl. oben).

Nach dem erfindungsgemäßen Verfahren erhältliche Nahrungsmittel enthalten also bereits nach der Herstellung große Mengen an Carotinoiden, die nicht zugesetzt werden müssen. Dadurch, dass das Nahrungsmittel neben dem mindestens einen Carotinoid auch Biomasse enthält ist sein Nährstoffgehalt zudem gesteigert. Insbesondere ist der Nährstoffgehalt nach der bevorzugten Alternative stark gesteigert, im dem es neben dem mindestens einen Carotinoid und Biomasse zusätzlich alle Medienbestandteile der Fermentation enthält. Es fallen demnach praktisch keine

Abfälle an, abgesehen von wässrigen Medien, welche jedoch unproblematisch in einer Kläranlage gereinigt werden können. Zusätzlich wird die gesamte Produktionsmenge Carotinoide ohne oder mit nur marginalen Verlusten verwertet, da keine verlustreichen Trenn- bzw. Aufarbeitungsschritte vorgenommen werden müssen, um die gesamte Menge Carotinoid zu extrahieren.

Die in dem oben beschriebenen, erfindungsgemäßen Verfahren eingesetzten Lösungsmittel werden alle soweit wie möglich aufbereitet und anschließend wieder verwendet bzw. dem Verfahren erneut zugeführt. Insbesondere wird das eingesetzte Dichlormethan bereits beim Lösungmittelaustausch gereinigt und steht anschließend zur erneuten Verwendung bereit. Der niedere Alkohol bzw. das Methanol wird z. B. destillativ gereinigt und ebenfalls wieder verwendet. Als Abfall fallen lediglich der Destillationssumpf an, der zusammen mit den wässrigen Medien gefahrlos einer Kläranlage zugeführt werden kann, wo letztendlich nur eine geringe Menge Klärschlamm als tatsächlicher Abfall anfällt. Somit ist das beschriebene Verfahren im wesentlichen abfallfrei.

20

30

15

10

Die Erfindung wird nachfolgend an Hand von Beispielen näher ausgeführt.

25 A) Kultivierung von Blakeslea trispora

Folgende Medien wurden zur Fermentation von Blakeslea trispora zur Produktion der Carotinoide eingesetzt:

Medium 1:

 Glucose
 10,00 g/l

 Baumwollsaatoel
 30,00 g/l

 Sojaoel
 30,00 g/l

	Dextrin	60,00 _, g/l
	Baumwollsamenmehl	75,00 g/l
	Triton X 100	1,20 g/l
	Ascorbinsäure	6,00 g/l
5	Milchsäure	2,00 g/l
	KH₂PO₄	0,50 g/l
	MnSO₄ x H2O	100 mg/l
	Thiamin-HCI	2 mg/l
	Isoniazid (Isonicotinsäurehydrazid)	0,75 g/l
10	Der pH wurde auf 6,5 eingestellt.	
	Medium 2:	
	Glucose	20 g/l
	Asparagin	2,00 g/l
15	KH ₂ PO ₄	5,00 g/l
	MgSO ₄ x 7 H₂O	0,50 g/l
	CaCl ₂	28 mg/l
	Thiamin-HCI	1,00 mg/l
	Citronensäure	2,00 mg/l
20	Fe(NO ₃) ₃ x 9 H ₂ O	1,50 mg/l
	ZnSO ₄ x 7 H₂O	1,00 mg/l
	MnSO ₄ x H ₂ O	0,30 mg/l
	CuSO ₄ x 5 H ₂ O	0,05 mg/l
	Na ₂ MoO ₄ x 2 H ₂ O	0,05 mg/l
25		
	•	
	Medium 3	
	Glucose	70,00 g/l
30	Asparagin	2,00 g/l
•	Hefe Extrakt	1,00 g/l

KH ₂ PO ₄	1,50 g/l
MgSO ₄ x 7 H ₂ O	0,50 g/l
Span 20	1,00 g/l
Thiamin-HCI	5,0 mg/l

5 Der pH wurde auf 5,5 eingestellt.

Mit Sporensuspensionen von Blakeslea trispora ATCC 14272 Mating Type (–) die 10⁸ (für Medium 2) bzw. 10⁷ (für Medium 1 und 3) Sporen enthielten, wurden je 200 ml der beschriebenen Medien angeimpft. Die Kultivierung erfolgte jeweils in 1-l-Erlenmeyerkolben mit Schikanen. Mit jedem Medium wurden sechs identische Kolben angesetzt und über 7 Tage bei 28°C und 140 UpM im Schüttler inkubiert.

B) Gentechnische Veränderung von Blakeslea Trispora

15

10

Material und Methoden

Molekulargenetische Arbeiten wurden, wenn nicht anders beschrieben, nach den Methoden in Current Protocols in Molecular Biology (Ausubel et al., 1999, John Wiley & Sons) durchgeführt.

20

25

30

Stämme und Wachstumsbedingungen

Die Blakeslea trispora Stämme ATCC 14271 (Paarungstyp(+)) und ATCC14272 (-) (ein Wildtyp) wurde erhaltenPaarungstyp (-)) wurden von der American Type Culture Collection. erhalten. Die Anzucht von B. trispora erfolgte in MEP-Medium (Malzextrakt-Pepton-Medium): 30 g/l Malzextrakt (Difco), 3 g/l Pepton (Soytone, Difco), 20 g/l Agar, Einstellung pH 5,5, ad 1000 ml mit H₂O bei 28 °C.

Die Anzucht von Agrobacterium tumefaciens LBA4404 erfolgte nach Hoekema et al. (1983, Nature 303:179-180) bei 28 °C für 24 h in Agrobacterien-Minimal Medium (AMM): 10 mM K₂HPO₄, 10 mM KH₂PO₄, 10 mM Glu-

20

25

cose, MM-Salze (2,5 mM NaCl, 2 mM MgSO₄, 700 µM CaCl₂, 9 µM Fe-SO₄, 4 mM (NH₄)₂SO₄).

Transformation von Agrobacterium tumefaciens

Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 5 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das A. tumefaciens Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den 10 binären Vektor).

Transformation von Blakeslea trispora

Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD600 von 0,15 in Induktionsmedium (IM: MM-Salze, 40 mM MES (pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 μM Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₀₀ von ca. 0,6 angezogen.

Zur Co-Inkubation von Blakeslea ATCC 14271 bzw. ATCC14272 und Agrobacterium wurden 100 µl Agrobakteriensuspension mit 100 µl Blakeslea Sporensuspension (107 Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt. Zur Selektion auf transformierte Blakesleazellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 30 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L- Asparagin, 50 mg/l MgSO₄ x 7H₂O, 150 mg/l KH₂PO₄, 25 μg/l ThiaminHCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, 100 mg/L Hygromycin, 100 mg/L Cefotaxim, pH 5,5,18 g/l Agar) ausplattiert. Zur Isolierung einzelner gentechnisch veränderter Sporen wurden die Sporen durch ein FACS Gerät der Fa. BectonDickson (Modell Vantage+Diva Option) einzeln auf Selektivmedium abgelegt.

Mutagenese mit MNNG

Zur Reduzierung der Anzahl von Kernen pro Spore wurde eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt. Hierfür wurde zunächst eine Sporensuspension mit 1 x 10⁷ Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 μg/ml zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.

Selektion homonukleater Zellen

Die Selektion homonukleater Zellen von Blakeslea trispora carB⁻ erfolgte analog zum Versuchsprotokoll für Phycomyces blakesleeanus (Roncero et al., 1984, Mutation Research, 125:195-204), modifiziert durch Wachstum in Gegenwart von 5-Carbon-5-Deazariboflavin (1 μg/ml) und Hygromycin 100 (μg/ml).

25

10

15

Herstellung genetisch veränderter Blakeslea trispora durch Agrobacterium-vermittelte Transformation

15

Herstellung des rekombinanten Plasmids pBinAHyg

Aus dem Plasmid pANsCos1 (Fig.1, Osiewacz, 1994, Curr. Genet. 26:87-90, SEQ ID NO: 4) wurde die gpdA-hph-trpC-Kassette als BgIII/HindIII Fragment isoliert und in das mit BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet (Fig. 2, SEQ ID NO: 3) und enthielt das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors (SEQ ID NO: 1) und des trpC Terminators (SEQ ID NO: 2) aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind. Die in den weiter unten beschriebenen Ausführungsbeispielen genannten Vektoren sind Abkömmlinge von pBinAHyg.

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in Agrobacterium tumefaciens

Nachfolgend wird beispielhaft die Übertragung des Plasmids pBinAHyg in Agrobacterien beschrieben. Die Übertragung der Abkömmlinge erfolgte analog.

Das Plasmid pBinAHyg wurde in den Agrobakterienstamm LBA 4404 (Hoekema et al., 1983, Nature 303:179-180) elektroporiert (Mozo and Hooykaas, 1991, Plant Mol. Biol. 16:917-918). Zur Selektion wurden bei der Agrobakterienanzucht folgende Antibiotika verwendet: Rifampicin 50 mg/l (Selektion auf das A. tumefaciens Chromosom), Streptomycin 30 mg/l (Selektion auf das Helferplasmid) und Kanamycin 100 mg/l (Selektion auf den binären Vektor).

Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in Blakeslea trispora

Zur Transformation wurden die Agrobakterien nach 24 h Anzucht in AMM auf eine OD₆₆₀ von 0,15 in Induktionsmedium (IM: MM-Saize, 40 mM MES

20

25

(pH 5,6), 5 mM Glucose, 2 mM Phosphat, 0,5% Glycerol, 200 μ M Acetosyringone) verdünnt und erneut über Nacht in IM bis zu einer OD₆₆₀ von ca. 0,6 angezogen.

Zur Co-Inkubation von Blakeslea trispora (B.t.) und Agrobacterium tume-faciens (A.t.) wurden 100 μl Agrobakteriensuspension mit 100 μl Blakeslea Sporensuspension (10⁷ Sporen/ml in 0,9% NaCl) gemischt und steril auf einer Nylon Membran (Hybond N, Amersham) auf IM-Agarose Platten (IM + 18 g/l Agar) verteilt. Nach 3 Tagen Inkubation bei 26 °C wurde die Membran auf eine MEP-Agarplatte (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) überführt.

Zur Selektion auf transformierte Blakeslea-Zellen enthielt das Medium Hygromycin in einer Konzentration von 100 mg/l sowie zur Selektion gegen Agrobakterien 100 mg/l Cefotaxim. Die Inkubation erfolgte für ca. 7 Tage bei 26 °C. Anschließend erfolgte der Transfer von Mycel auf frische Selektionsplatten. Gebildete Sporen wurden mit 0,9% NaCl abgespült und auf CM17-1-Agar (3 g/l Glucose, 200 mg/l L-Asparagin, 50 mg/l MgSO₄ x 7H₂O, 150 mg/l KH2PO4, 25 µg/l Thiamin-HCl, 100 mg/l Yeast Extract, 100 mg/l Na-desoxycholat, pH 5,5, 100 mg/l Cefotaxim, 100 mg/l Hygromycin, 18 g/l Agar) ausplattiert. Die Übertragung von Sporen auf frische Selektionsplatten wurde dreimal wiederholt. Auf diese Weise wurde die Transformante Blakeslea trispora GVO 3005 isoliert. Alternativ erfolgte zur Selektion der GVO (gentechnisch veränderten Organismen) die Einzelablage der Sporen durch den BectonDickinson FacsVantage+Diva Option auf CM-17 Agar mit 100 mg/l Cefotaxim, 100 mg/l Hygromycin. In diesem Fall wurde nur dort Pilzmycel gebildet, wo die Sporen gentechnisch verändert waren.

Nachweis der genetischen Veränderung durch Übertragung von pBinAHyg und Abkömmlingen von pBinAHyg in Blakeslea trispora Nachfolgend wird beispielhaft der Nachweis der Übertragung für pBinA-Hyg in Blakeslea trispora beschrieben. Der Nachweis der Übertragung der Abkömmlinge erfolgte analog.

5 200 ml MEP-Medium (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5) wurden mit 10⁵ bis 10⁷ Sporen der Transformante Blakeslea trispora GVO 3005 beimpft und 7 Tage bei 26 °C mit 200 Upm auf einem Rundschüttler inkubiert. Zum Nachweis der erfolgreichen Transformation wurde DNA aus dem Mycel isoliert (Peqlab Fungal DNA Mini Kit) und in einer PCR (Programm: 94 °C 1 min, dann 30 Zyklen mit 1 min. 94°C, 1 min. 58 °C, 1 min. 72 °C) eingesetzt.

Zum Nachweis des Hygromycinresistenzgens (hph) wurden die Primer hph-forward (5'-CGATGTAGGAGGGCGTGGATA, SEQ ID NO: 5) und hph-reverse (5'-GCTTCTGCGGGCGATTTGTGT, SEQ ID NO: 6) verwendet. Das erwartete Fragment von hph wies eine Länge von 800 bp auf.

Zur Amplifikation des Kanamycinresistenzgens nptlll und damit als Kontrolle auf Agrobakterien wurden die Primer nptlll-forward (5'-TGAGAATATCACCGGAATTG, SEQ ID NO: 7) und nptlll-reverse (5'-AGCTCGACATACTGTTCTTCC, SEQ ID NO: 8) verwendet. Das erwartete Fragment von nptlll wies eine Länge von 700 bp auf.

Glycerinaldehyd-3eines Fragmentes des Amplifikation Zur phosphatdehydrogenasegens gpd1 und damit als Kontrolle auf Blakeslea (5'-Primer MAT292 wurden die trispora GTGAATGGAAATCCCATCGCTGTC, SEQ ID NO: 9) und MAT293 (5'-AGTGGGTACTCTAAAGGCCATACC, SEQ ID NO: 10) verwendet. Das erwartete Fragment von gpd1 wies eine Länge von 500 bp auf.

25

15

20

Das Ergebnis der PCR der Blakeslea trispora DNA ist in Fig. 3 anhand eines Standard-Gels gezeigt. Die Spuren des Gels wurden folgendermaßen belegt:

5 1) 100 bp Größenmarker (100 bp - 1 kb)

2) B.t. GVO 3005 primer nptlll-for / nptlll-rev
3) B.t. GVO 3005 primer hph-for / hph-rev
4) B.t. GVO 3005 primer MAT292 / MAT293 (gpd)
5) A.t. mit Plasmid pBinAHyg primer nptlll-for / nptlll-rev

10 6) A.t. mit Plasmid pBinAHyg primer hph-for / hph-rev

7) B.t. 14272 WT primer nptlll-for / nptlll-rev
8) B.t. 14272 WT primer hph-for / hph-rev

9) B.t. 14272 WT primer MAT292 / MAT293 (gpd)

In der DNA von Blakeslea trispora wurde das Hygromycinresistenzgens (hph) und als Positivkontrolle Glycerinaldehyd-3-phosphatdehydrogenasegen (gpd1) nachgewiesen. nptlll konnte demgegenüber nicht nachgewiesen werden.

20 Somit wurde die genetische Veränderung von Blakeslea trispora durch Agrobacterium-vermittelte Transformation nachgewiesen.

Isolierung homokaryotischer GVO von Blakeslea trispora:

25 Herstellung homonukleater Stämme

30

Durch erfolgreichen Transfer des Vectors pBinAHyg und Abkömmlingen von pBinAHyg in Blakeslea trispora entstandenentstehen genetisch veränderte Organismen. In GVO von Blakeslea trispora. Jedoch liegen in Blakeslea liegen in allen Stadien des vegetativen und des sexuellen Zellzyklus mehrkernige Zellen vor. Daher erfolgte erfolgt die Insertion der VectorFremd-DNA in der Regel nur in einem Kern. Ziel ist es aber, dass,

20

25

30

Stämme von Blakeslea zu erhalten, bei denen die Insertion der Vector-Fremd-DNA in allen Kernen vorliegt., d.h. Ziel ist ein homonukleates rekombinantes Pilzmycel.

Zur Herstellung solcher homokaryotischer Zellen wurden zunächst Sporensuspensionen der rekombinanten Stämme mit MNNG behandelt. Hierfür wurde eine Sporensuspension mit 1 x 10⁷ Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 μg/ml zugegeben. Die Dauer der Inkubation mit MNNG wurde so gewählt, dass die Überlebensrate der Sporen ~5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und plattiert.

1) Herstellung homonukleater rekombinanter Stämme durch FACS (fluorescence-activated cell sorting)

Ein geringer Anteil der Sporen von Blakeslea trispora bzw. der gentechnisch veränderten Stämme von Blakeslea trispora ist von Natur aus einkernig. Zur Herstellung homonukleater rekombinanter Stämme, die Fremd-DNA von pBinAHyg oder pBinAHyg-Abkömmlingen enthielten, wurden die einkernigen Sporen durch FACS aussortiert und auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die hier gebildten Mycelien waren homonukleat. Zur Sortierung mit FACS wurden die Sporen eines 3 Tage alten Ausstriches mit 10 ml Tris-HCl 50mMol + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis 0,8 x 10⁷ Sporen pro ml. Zu 9 ml Sporensuspension wurden 1ml DMSO und 10 µl Syto 11 (Farbstoff-Stammlösung in DMSO Molecular Probes Nr.S-7573) zugegeben. Danach wurde 2 h bei 30°C gefärbt. Die Selektion und Ablage erfolgte mittels eines BectonDickinson FacsVantage+Diva Option. Die Selektion erfolgt zuerst nach Größe, um einzelne Sporen von Aggregaten und Verunreinigungen zu trennen. Dann wurden diese Sporen nach ihrer Fluoreszenz (Anregung = 488nm Emission = 530 nm) sortiert abgelegt. Die linke Schulter der Gauß-Kurve der Fluoreszenzhäufigkeitsverteilung enthielt die einkernigen Sporen.

5 Anschließend wurden die Sporen auf MEP-Agarplatten ausplattiert und neue Sporen erzeugt.

Diese Sporen wurden analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin plattiert, das zusätzlich Hygromycin enthielt.

Hierdurch wurden homokaryonte Zellen des Genotyps hyg^R und dar selektiert.

15

20

25

10

2) Herstellung homonukleater Stämme durch Kernreduktion und Selektion mit FACS

Zur Reduzierung der Anzahl von Kernen pro Spore wurde vor der Selektion eine Behandlung von Sporensuspensionen mit MNNG (N-Methyl-N'-nitro-N-nitrosoguanidin) durchgeführt, und so durch chemische Mutagenese eine Kernreduktion erzielt.

Hierfür wurde zunächst eine Sporensuspension mit 1 x 10⁷ Sporen/ml in Tris/HCl-Puffer, pH 7,0 hergestellt. Der Sporensuspension wurde MNNG in einer Endkonzentration von 100 µg/ml zugegeben. Die Zeit der Inkubation in MNNG wurde so gewählt, dass die Überlebensrate der Sporen ca. 5% betrug. Nach Inkubation mit MNNG wurden die Sporen dreimal mit 1g/l Span 20 in 50 mM Phosphatpuffer pH 7,0 gewaschen und nach der unter 1) beschriebenen Methode sortiert bzw. selektiert.

Alternativ konnten zur Reduktion der Kernzahl in den Sporen auch Röntgen – und UV-Strahlen eingesetzt werden, wie es von Cerdá-Olmedo und Patricia Reau in Mutation Res., 9 (1970), 369-384 beschrieben wurde.

5

10

15

3) Herstellung homonukleater Stämme durch Selektion auf rezessive Selektionsmarker

Als rezessiver Selektionsmarker zur Selektion homonukleater Mycelien kommt beispielsweise der rezessive Selektionsmarker pyrG in Frage. Wildtyp-Stämme von Blakeslea trispora sind pyrG⁺. Diese Stämme können nicht in Gegenwart des Pyrimidin-Analogs 5-Fluororotat (FOA) wachsen, weil sie FOA durch die Orotidin-5'-monophosphatdecarboxylase zu lethalen Metaboliten umsetzen. Gentechnisch veränderte Blakesleaa, die ho-Orotidin-5'-Enzymaktivität fehlt die pvrG⁻ sind, monukleat monophosphatdecarboxylase. Folglich können diese pyrG-Stämme 5-Fluororotat nicht verwerten. Die Stämme wachsen daher in Gegenwart von FOA und Uracil. Im Fall der Kopplung der Mutation pyrG- und der Insertion von Fremd-DNA auf dem Kern einer einkernigen Spore, kann aus dieser Spore homonukleates rekombinantes Pilzmycel gebildet werden.

20

25

Zunächst wurde durch Insertion eines Fragentes von pyrG (SEQ ID NO: 65) aus Blakeslea trispora in pBinAHyg das Plasmid pBinAHygBTpyrG-SCO (SEQ ID NO: 36, Fig. 4) erzeugt. Dieses Plasmid wurde in Blakelea trispora transformiert und führte dort durch homologe Rekombination zur Disruption von pyrG.

30

Homonukleate GVO von Blakeslea trispora mit dem Phänotyp pyrG⁻ wurden folgendermaßen selektiert. Zur Agrobakterium-vermittelten Transformation von pBinAHygBTpyrG-SCO wurde wie oben beschrieben auf MEP (30 g/l Malzextrakt, 3 g/l Pepton, pH 5,5, 18 g/l Agar) mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin plattiert. Die Sporen der Transformanten

wurden mit 10 ml Tris-HCl 50mM + 0,1% Span20 pro Agar-Platte abgeschwemmt. Die Sporenkonzentration betrug 0,5 bis 0,8 x 10⁷ Sporen pro ml. Die Sporen wurden anschließend auf FOA-Medium mit 100 mg/l Cefotaxim und 100 mg/l Hygromycin ausplattiert. FOA-Medium enthielt pro Liter 20 g Glucose, 1 g FOA, 50 mg Uracil, 200 ml Citrat-Puffer (0,5 M, pH 4,5) und 40 ml Spurensalzlösung nach Sutter, 1975, PNAS, 72:127). Homonukleate pyrG⁻-Mutanten zeigten Wachstum auf dem Uracil-haltigen FOA-Medium; aber kein Wachstum bei Plattierung auf FOA-Medium ohne Uracil. Auf die gleiche Weise wurden aus den im folgenden beschriebenen GVO von Blakeslea trispora zur Herstellung von Xanthophyllen homonukleate GVO hergestellt.

Alternativ ist es möglich die Sporen analog zur Vorschrift von Roncero et al. auf Medium mit 5-Carbon-5-deazariboflavin zu plattieren, das zusätzlich Hygromycin enthält (Roncero et al., 1984, Mutation Research, 125: 195 - 204). Hierdurch werden homokaryonte Zellen des Genotyps hyg^R und dar selektiert. Nach diesem Prinzip werden homokaryonte Stämme von Blakeslea trispora mit dem Phänotyp hyg^R und dar erzeugt.

20

25

30

15

10

Ausführungsbeispiele zur Herstellung von gentechnisch veränderten Organismen von Blakeslea trispora für die Herstellung von Carotinoiden und Carotinoidvorstufen

Die Erzeugung der im folgenden genannten Plasmide erfolgte durch die Methode "overlap-extension PCR" und durch anschließende Insertion der Amplifikationsprodukte in das Plasmid pBinAHyg. Die Methode "overlapextension PCR" erfolgte wie in Innis et al. (Eds.) PCR protocols: a guide to methods and applications, Academic Press, San Diego beschrieben. Die Transformation der pBinAHyg-Abkömmlinge und die Herstellung homonukleater gentechnisch veränderter Stämme von Blakeslea trispora erfolgte wie oben beschrieben.

30

Gentechnisch veränderte Stämme von Blakeslea trispora zur Herstellung von Zeaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von Blakeslea trispora für die Herstellung von Zeaxanthin verwendet, codieren also u.a. Hydroxylasen (crtZ):

- p-tef1-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ (SEQ ID NO: 70) aus Haematococcus pluvialis Flotow NIES-144 (Accession No. AF162276) unter Kontrolle des ptef1 Promotors aus Blakeslea trispora (Seq. pBinAHygBTpTEF1-HPcrtZ, SEQ ID NO: 37, Fig. 5);
- p-carRA-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus Blakeslea trispora (Seq. pBinAHyg-BTpcarRA-HPcrtZ, SEQ ID NO: 38, Fig. 6)
- p-carB-HPcrtZ, enthaltend Gen der Hydroxylase HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 unter Kontrolle des Promotors pcarB aus Blakeslea trispora (Seq. pBinAHygBTpcarB-HPcrtZ, SEQ ID NO: 39, Fig. 7)
- p-carRA-HPcrtZ-TAG-3'carA-IR, enthaltend Gen der Hydroxylase
 HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus Blakeslea trispora. Stromabwärts des Gens der Hydroxylase ist eine Inverted-Repeat-Struktur lokalisiert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp "Loop" und anschließend ca. 350 bp "Inverted Repeat 2") (Seq. pBinAHyg-BTpcarRA-HPcrtZ-TAG-3'carA-IR, SEQ ID NO: 40, Fig. 8);
 - p-carRA-HPcrtZ-GCG-3'carA-IR, enthaltend Gen der Hydroxylase
 HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 unter Kontrolle des Promotors pcarRA aus Blakeslea trispora. Das Gen der

20

25

Hydroxylase ist mit einer Inverted-Repeat-Struktur fusioniert, die aus dem 3'-Ende von carA und der stromabwärts von carA gelegenen Region stammt (IR, SEQ ID NO: 74, "Inverted Repeat 1' ca. 350 bp von carA, dann ca. 200 bp "Loop" und anschließend ca. 350 bp "Inverted Repeat 2"). Das abgeleitete Fusionsprotein besteht folglich aus der Hydroxylase von Haematococcus pluvialis und dem Carboxyterminus von CarA aus Blakeslea trispora (Seq. pBinAHyg-BTpcarRA-HPcrtZ-GCG-3'carA-IR, SEQ ID NO: 41, Fig. 9);

- p-tef1-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ (SEQ ID NO: 71) aus Erwinia uredova 20D3 (Accession No. D90087) unter Kontrolle des ptef1 Promotors (Seq. pBinAHygBTpTEF1-EUcrtZ, SEQ ID NO: 42, Fig. 10);
- p-carRA-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus Erwinia uredova 20D3 unter Kontrolle des Promotors pcarRA aus Blakeslea trispora (Seq. pBinAHygBTpcarRA-EUcrtZ, SEQ ID NO: 43, Fig. 11);
 - p-carB-EUcrtZ, enthaltend Gen der Hydroxylase EUcrtZ aus Erwinia uredova 20D3 unter Kontrolle des Promotors pcarB aus Blakeslea trispora (Seq. pBinAHygBTpcarB-EUcrtZ, SEQ ID NO: 44, Fig. 12);
 - p-gpdA-HPcrtZ-t-crtZ, enthaltend Gen der Hydroxylase HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 unter Kontrolle des gpdA Promotors und des Terminators t-crtZ; d.h. des stromabwärts von crtZ aus Haematococcus pluvialis Flotow NIES-144 gelegenen Sequenzabschnitts (SEQ ID NO: 73) (Seq. pBinAHyg-gpdA-HPcrtZtcrtZ, SEQ ID NO: 45, Fig. 13).
 - p-gpdA-BTcarR-HPcrtZ-BTcarA, enthaltend Genfusion aus Genen der Lycopincyclase carR aus Blakeslea trispora, der Hydroxylase HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 und der

Phytoensynthase carA aus Blakeslea trispora unter Kontrolle des gpdA Promotors aus Aspergillus nidulans (Seq. pBinAHyg-carR_crtZ_carA, SEQ ID NO: 46, Fig. 14);

5 Herstellung gentechnisch veränderter Stämme von Blakeslea trispora zur Herstellung von Canthaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von Blakeslea trispora für die Herstellung von Canthaxanthin verwendet, codieren also u.a. Ketolasen (crtW):

- p-tef1-NPcrtW, enthaltend das Gen der Ketolase NPcrtW (SEQ ID NO: 72) aus Nostoc punctiforme PCC73102 (ORF148, Accesion No. NZ_AABC01000196) unter Kontrolle des ptef1 Promotors aus Blakeslea trispora (Seq. pBinAHygBTpTEF1-NpucrtW, SEQ ID NO: 47, Fig. 15);
- p-carRA-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus Nostoc punctiforme PCC73102 unter der Kontrolle des Promotors pcarRA aus Blakeslea trispora (Seq. pBinAHygBTpcarRA-NpucrtW, SEQ ID NO: 48, Fig. 16);
- p-carB-NPcrtW, enthaltend das Gen der Ketolase NPcrtW aus No stoc punctiforme PCC73102 unter der Kontrolle des Promotors pcarB aus Blakeslea trispora (Seq. pBinAHygBTpcarB-NpucrtW, SEQ ID NO: 49, Fig. 17);

Herstellung gentechnisch veränderter Stämme von Blakeslea trispora zur Herstellung von Astaxanthin

Folgende Plasmide (Abkömmlinge von pBinAHyg) wurden zur gentechnischen Veränderung von Blakeslea trispora für die Herstellung von Astaxanthin verwendet, codieren also u.a. für Hydroxylasen (crtZ) und Ketolasen (crtW):

- p-carRA-HPcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxylase HPcrtZ aus Haematococcus pluvialis Flotow NIES-144 und das Gen der Ketolase NPcrtW aus Nostoc punctiforme PCC73102 (ORF148, Accesion No. NZ_AABC01000196) beide jeweils unter Kontrolle des Promotors pcarRA aus Blakeslea trispora (Seq. pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 50, Fig. 18);
- p-carRA-EUcrtZ-pcarRA-NPcrtW, enthaltend das Gen der Hydroxy-lase EUcrtZ aus Erwinia uredova20D3 (Accession No. D90087) und
 das Gen der Ketolase NPcrtW aus Nostoc punctiforme PCC73102 beide jeweils unter Kontrolle des Promotors pcarRA aus Blakeslea trispora (Seq. pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW, SEQ ID NO: 51, Fig. 19);
- 15 Klonierung und Sequenzanalyse von Genen und Promotoren, die beispielhaft für die gentechnische Veränderung von Blakeslea trispora genutzt werden können.

Nachfolgend werden beispielhaft die Klonierung und Sequenzierung verschiedener Gene und Promotoren aus Blakeslea trispora beschrieben.

20

25

30

Klonierung und Sequenzanalyse ptef1

Die Klonierung von p-tef aus Blakeslea trispora erfolgte auf der Grundlage einer bereits in GenBank veröffentlichten Sequenz des Strukturgens für den Translations-Elongationsfaktor 1-α aus Blakeslea trispora (AF157235). Ausgehend von dem Sequenzeintrag AF157235 wurden Primer für die inverse PCR ausgewählt, um die stromaufwärts des Strukturgens gelegene Promotoregion zu amplifizieren und zu sequenzieren. In der inversen nested PCR an 200 ng Xhol-gespaltener und zirkularisierter genomischer DNA von Blakeslea trispora ATCC14272 wurde ein 3000-bp-Fragment in folgendem Ansatz erhalten: Matrizen-DNA (1 μg genomi-

sche DNA von Blakeslea trispora ATCC 14272) Primer MAT344 5'-GGCGTACTTGAAGGAACCCTTACCG-3' (SEQ ID NO: 63) und MAT 345 5'-ATTGATGCTCCCGGTCACCGTGATT-3' (SEQ ID NO: 64) je 0,25 μM, 100 μM dNTP, 10 μl Herculase-Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H₂O ad 100 μl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus). Der Sequenzabschnitt, der stromaufwärts des vermutlichen Startcodons des Gens tef1 innerhalb 3000-bp-Fragmentes liegt, wurde als Promotor ptef1 bezeichnet.

10

15

20

25

30

Zyklen); 72 °C, 10 min (1 Zyklus).

Klonierung Sequenzanalyse des Gens der HMG-CoA-Reduktase aus Blakeslea trispora

Zunächst wurde mit dem Cosmidvektor pANsCos1 eine Genbank von Blakeslea trispora ATCC 14272, Mating Type (-) hergestellt. Der Vektor wurde durch Spaltung mit Xbal linearisiert und anschließend dephosphoryliert. Eine weitere Spaltung mit mit BamHI schuf die Insertionsstelle, in welche die mit Sau3Al partiell gespaltene und dephosphorylierte genomische DNA von Blakeslea trispora ligiert wurde. Die derart gebildeten Cosmide wurden anschließend in vitro verpackt und in Escherichia coli übertragen. Auf der Grundlage der bekannten Sequenz eines Fragmentes des HMG-CoA-Reduktase codierenden Gens aus Blakeslea trispora (Eur. J. Biochem 220, 403-408 (1994)) wurde eine 315-bp-DNA-Sonde durch folgende PCR hergestellt. Reaktionsansatz: 1 µg genomische DNA von Blakes-**MAT314** 5'lea ATCC 14272, Primer trispora CCGATGGCGACGACGGAAGGTTGTT-3' [SEQ ID NO 79] und MAT315 5'-CATGTTCATGCCCATTGCATCACCT-3' [SEQ ID NO 80] je 0,25 μM, 100 μM dNTP, 10 μl Herculase-Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H2O ad 100 µl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 58 °C, 30 s. 72 °C, 30 s, 95 °C, 30 s (30

10

15

25

Mit dieser DNA-Sonde wurde die Cosmid-Genbank durchmustert. Es wurde ein Klon identifiziert, dessen Cosmid mit der DNA-Sonde hybridisierte. Die Insertion dieses Cosmids wurde sequenziert. Die DNA-Sequenz enthielt einen Abschnitt, der dem Gen einer HMG-CoA-Reduktase zugeordnet wurde [HMG-CoA-Red.gb].

Klonierung und Sequenzanalyse carB

(carB = Gen der Phytoendesaturase aus Blakeslea trispora)

Aus dem Sequenzvergleich der Peptidsequenzen von Phytoendesaturasen und dem Vergleich der zugehörigen DNA-Sequenzen von Phycomyces blakesleeanus, Cercospora nicotianae, Phaffia rhodozyma und Neurospora crassa wurden die degenerierten Primer MAT182 5'-5'-(SEQ 52) **MAT192** GCNGARGGNATHTGGTA-3' ID und TCNGCNAGRAADATRTTRTG-3 (SEQ ID 53) abgeleitet. Die PCR wurde in 100 ul Ansätzen durchgeführt. Diese enthielten 200 ng genomische DNA von Blakeslea trispora ATCC14272, 1 µM MAT182, 1 µM MAT192, 100 µM dNTP, 10 µl Pfu-Polymerasepuffer 10x, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C), H₂O ad 100 µl.

Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 40
 °C, 30 s, 72 °C, 30 s, 95 °C, 30 s (35 Zyklen); 72 °C, 10 min (1 Zyklus).

Hiermit wurde ein 358-bp-Fragment erhalten, dessen abgeleitete Peptidsequenz Ähnlichkeit zu den Sequenzen der Phytoendesaturasen aufwies. Durch die Methode der inversen PCR (Innis et al. in PCR protocols: a guide to methods and applications. 1990. S. 219-227) wurden nach dem Prinzip des Chromosome-Walking die Genregionen stromaufwärts und stromabwärts des 350-bp-Fragmentes folgendermaßen amplifiziert, kloniert und sequenziert:

30 (i) ein 1,1-kbp-Fragment durch PCR mit den Primern MAT219 5'-AAGTGACACCGGTTACACGCTTGTCTT-3' (SEQ ID 54) und MAT

20

25

220 5'-GCTTATCACCATCTGTTACCTCCTTGC-3' (SEQ ID 55) erhalten aus 200 ng EcoRI-gespaltener und zirkularisierter genomischer DNA von Blakeslea trispora ATCC14272, 0,25 μM MAT219, 0,25 μM MAT220, 100 μM dNTP, 10 μI Herculase-Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H_2O ad 100 μl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s. 72 °C, 60 s, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus),

ein 2,9-kbp-Fragment durch PCR mit den Primern MAT219 und MAT220 erhalten aus 200 ng Xbal-gespaltener und zirkularisierter genomischer DNA von Blakeslea trispora ATCC14272, 0,25 μΜ MAT219, 0,25 μΜ MAT220, 100 μΜ dNTP, 10 μl Herculase-Polymerasepuffer 10x, 5 U Herculase (Zugabe bei 85 °C), H₂O ad 100 μl. Das PCR-Profil war 95 °C, 10 min (1 Zyklus); 85 °C, 5 min (1 Zyklus); 60 °C, 30 s, 72 °C, 3 min, 95 °C, 30 s (30 Zyklen); 72 °C, 10 min (1 Zyklus);

Der klonierte Sequenzabschnitt ist schematisch in Fig. 20 (SEQ ID NO 77) dargestellt. Die Sequenzierung erfolgte in Strang- und Gegenstrangrichtung mit den klonierten Fragmenten sowie mit den PCR-Produkten. Die Sequenz des klonierten Sequenzabschnitts ist in Fig. 21 (SEQ ID NO 78) gezeigt.

Sequenzvergleiche

Die Nukleotidsequenz von carB und die Peptidsequenz des abgeleiteten Proteins CarB wurden mit den bekannten Sequenzen verwandter Proteine verglichen. Zum Sequenzvergleich wurden die Programme GAP und BESTFIT eingesetzt.

30 CarB - Identische Aminoacylreste nach GAP

Programmeinstellungen:

PCT/EP2004/000099 WO 2004/063359

78

Gap Weight: 8

2 Length Weight:

2.912

Average Match:

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu 5

CarB aus Blakeslea trispora ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 72,491

Phaffia rhodozyma:

50,460

Neurospora crassa:

47,943

Cercospora nicotianae: 10

47,740

CarB -Identische Aminoacylreste nach BESTFIT

Programmeinstellungen:

Gap Weight: 8

Length Weight: 15

2

Average Match:

2.912

Average Mismatch: -2.003

Dabei wurde folgende Werte für die Übereinstimmung der Aminosäuren zu

CarB aus Blakeslea trispora ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 73,380 20

Phaffia rhodozyma:

53,175

Neurospora crassa:

51,896

Cercospora nicotianae:

50,791

carB - Identische Basen nach GAP 25

Programmeinstellungen:

Gap Weight:

50

Length Weight:

3

Average Match:

10.000

Average Mismatch: 0.000 30

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB aus Blakeslea trispora ATCC14272 in % gefunden:

Phycomyces blakesleeanus: 64,853

Cercospora nicotianae:

50.143

5 Phaffia rhodozyma: 43,179

Neurospora crassa:

42,130

carB -Identische Basen nach BESTFIT

Programmeinstellungen:

Gap Weight: 10

50

Length Weight:

3

Average Match:

10.000

Average Mismatch: -9.000

Dabei wurde folgende Werte für die Übereinstimmung der Basen zu CarB

aus Blakeslea trispora ATCC14272 in % gefunden: 15

Phycomyces blakesleeanus: 68,926

Phaffia rhodozyma:

62,403

Neurospora crassa:

60,230

Cercospora nicotianae:

56,884

20

25

30

Klonierung zur Expression von carB

Zur Klonierung und Expression von carB aus Blakeslea trispora wurden von dem oben beschriebenen klonierten Sequenzabschnitt aus Blakeslea trispora in sechs Leserastern die möglichen Proteinsequenzen abgeleitet. Diese Proteinsequenzen wurden mit den Sequenzen der Phytoendesaturasen aus Phycomyces blakesleeanus, Phaffia rhodozyma, Neurospora crassa, Cercospora nicotianae verglichen. Auf der Grundlage des Sequenzvergleiches wurden im klonierten Sequenzabschnitt der genomischen DNA von Blakeslea trispora drei Exons identifiziert, die zusammengefügt eine codierende Region ergeben, deren abgeleitetes Genprodukt über die gesamte Länge 72,7% identische Aminoacylreste mit der Phy-

15

20

25

30

toendesaturase CarB aus Phycomyces blakesieeanus aufweist. Dieser Sequenzabschnitt aus drei möglichen Exons und zwei möglichen Introns wurde daher als Gen carB bezeichnet. Zur Überprüfung der vorhergesagten Genstruktur wurde die codierende Sequenz von carB aus Blakeslea trispora durch PCR mit cDNA von Blakeslea trispora als Matrize und mit 5'-Bol1425 Primern den AGAGAGGGATCCTTAAATGCGAATATCGTTGC-3' (SEQ ID 56) und Bol1426 5'-AGAGAGGGATCCATGTCTGATCAAAAGAAGCA-3' (SEQ ID 57) erzeugt. Das erhaltene DNA-Fragment wurde sequenziert. Die Lokalisation von Exons und Introns wurde durch Vergleich der cDNA mit der genomischen DNA von carB bestätigt. In Fig. 21 ist die codierende Sequenz von carB schematisch dargestellt. Zur Expression von carB in Escherichia coli wurde zunächst die Ndel-Schnittstelle in carB durch die Methode overlap extension PCR entfernt sowie am 5'-Ende des Gens eine Ndel-Schnittstelle und am 3'-Ende eine BamHl-Schnittstelle eingefügt. Das erhaltene DNA-Fragment wurde mit dem Vektor pJOE2702 ligiert. Das erhaltene Plasmid wurde als pBT4 bezeichnet und zusammen mit pCAR-AE in Escherichia coli XL1-Blue kloniert. Die Expression erfolgte durch Induktion mit Rhamnose. Der Nachweis der Enzymaktivität erfolgte durch Nachweis der Lycopinsynthese via HPLC. Die Klonierungsschritte sind im folgenden beschrieben:

PCR 1.1:

Temperaturprofil:

Ca. 0,5 μg cDNA von Blakeslea trispora, 0,25 μM MAT350 5'ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID 58),
0,25 μM MAT244 5'GTTCCAATTGGCCACATGAAGAGTAAGACAGGAAACAG-3' (SEQ ID
59), 100 μM dNTP, 10 μl Pfu-Polymerase-Puffer (I0x), 2,5 U PfuPolymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100μL.

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30 s, 9. 72 °C 10min Zyklen: (1-2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

5 PCR1.2:

10

15

Ca. 0,5 μ g cDNA von Blakeslea trispora, 0,25 μ M MAT243 5'-CCTGTCTTACTCTTCATGTGGCCAATTGGAACCAACAC-3' (SEQ ID 60), 0,25 μ M MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID 61), 100 μ M dNTP, 10 μ l Pfu-Polymerase-Puffer (I0x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H₂O ad 100 μ L.

Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 40 °C 30s, 4. 72 °C 1 min 30 s, 5. 95 °C 30 s, 6. 50 °C 30 s, 7. 72 °C 1 min 30 s, 8. 95 °C 30s, 9. 72 °C 10min Zyklen: (1 -2.) 1x, (3-5.) 5x, (6-8.) 25x, (9.) 1x

Reinigung der PCR-Fragmente aus PCR 1.1, 1.2

Dazu wurde PCR 2 zur Herstellung der codierenden Sequenz von carB aus Blakeslea trispora für die Klonierung in pJOE2702 durchgeführt:

- Ca. 50 ng Produkt aus PCR 1.1 und ca. 50 ng Produkt aus PCR1.2 mit 0,25 μ M MAT350 5'-ACTTTATTGGATCCTTAAATGCGAATATCGTTGCTGC-3' (SEQ ID NO 58), 0,25 μ M MAT353 5'-CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3' (SEQ ID NO
- 25 61), 100 μM dNTP, 10 μL Pfu-Polymerase-Puffer (l0x), 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und H_2O ad 100 μL.

Temperaturprofil:

- 1. 95°C 10 min, 2. 85 °C 5 min, 3. 59 °C 30 s, 4. 72 °C 2 min, 5. 95 °C 30 s, 6.72°C 10 min
- 30 Zyklen: (1-2.) 1x, (3-5.) 22x, (6.) 1x

15

Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Ligation in Vektor pPCR-Script-Amp, eine Klonierung in Escherichia coli XL1-Blue, Sequenzierung der Insertion, Spaltung mit Ndel und BamHI sowie eine Ligation in pJOE2702. Das erhaltene Plasmid wurde als pBT4 bezeichnet.

Charakterisierung und Nachweis der Enzymaktivität von CarB (Phytoendesaturase)

Das von carB abgeleitete Genprodukt wurde als CarB bezeichnet. CarB weist auf Grundlage der Peptidsequenzanalyse folgende Eigenschaften auf:

Länge: 582 Aminoacylreste

Molekulare Masse: 66470

Isoelektrische Punkt: 6,7

Katalytische Aktivität: Phytoendesaturase

Edukt: Phytoen
Produkt: Lycopin

EC-Nummer: EC 1.14.99-

Der Nachweis der Enzymaktivität erfolgte in vivo. Wenn das Plasmid (pCAR-AE) in Escherichia coli XL1-Blue übertragen wird, entsteht der Stamm Escherichia coli XL1-Blue (pCAR-AE). Dieser Stamm synthetisiert Phytoen. Wenn zusätzlich das Plasmid pBT4 in Escherichia coli XL1-Blue übertragen wird, entsteht der Stamm Escherichia coli XL1-Blue (pCAR-AE)(pBT4). Da ausgehend von carB eine enzymatisch aktive Phytoendesaturase gebildet wird, produziert dieser Stamm Lycopin.

Die Plasmide pCAR-AE und pBT4 wurden daher in Escherichia coli übertragen. Nach Wachstum in Flüssigkultur wurden die Carotinoide aus den Zellen extrahiert und charakterisiert (vgl. oben).

WO 2004/063359 PCT/EP2004/000099

Durch HPLC Analyse wurde nachgewiesen, daß der Stamm Escherichia coli XL1-Blue (pCAR-AE) Phytoen und der Stamm Escherichia coli XL1-Blue (pCAR-AE)(pBT4) Lycopin produziert. CarB weist folglich die Enzymaktivität einer Phytoendesaturase auf.

5

Herstellung gentechnisch veränderter Stämme von Blakeslea trispora zur Herstellung von Phytoen

Nachfolgend werden beispielhaft die Herstellung von gentechnisch veränderten Organismen zur Herstellung von Phytoen beschrieben.

10

15

20

25

Vector pBinAHyg∆carB zur Erzeugung von carB⁻ -Mutanten von Blakeslea trispora

Für die Deletion von carB in Blakeslea trispora wurde der Vektor pBinA-Hyg∆carB (SEQ. ID. NO:62, Fig. 22) konstruiert. Der Vorläufer von pBinAHyg∆carB ist pBinAHyg (SEQ. ID. NO:3, Fig. 2). pBinAHyg wurde folgendermaßen konstruiert:

Aus dem Plasmid pANsCos1 (SEQ. ID. NO:4, Fig. 1, Osiewacz, 1994, Curr. Genet. 26:87-90) wurde die gpdA-hph Kassette als BgIII/HindIII Fragment isoliert und in das BamHI/HindIII geöffnete binäre Plasmid pBin19 (Bevan, 1984, Nucleic Acids Res. 12:8711-8721) ligiert. Der so erhaltene Vektor wurde als pBinAHyg bezeichnet und enthält das *E. coli* Hygromycin-Resistenzgen (hph) unter Kontrolle des gpd Promotors und des trpC Terrminators aus *Aspergillus nidulans* sowie die entsprechenden Bordersequenzen, die für den DNA-Transfer von *Agrobacterium* notwendig sind.

Die Amplifikation der codierenden Sequenz von carB mit den Primern MAT350 (SEQ ID NO 58) und MAT353 (SEQ ID NO 61) mittels PCR wurde mit den folgenden Parametern durchgeführt:

30 50 ng pBT4 mit 0,25 μ M MAT350 5'-ACTTTATTGGATCCTTAAAT-GCGATCCTGC-3', 0,25 μ M MAT353 5'-

CTATTTTAATCATATGTCTGATCAAAAGAAGCATATTG-3', 100 μM dNTP, 10 µL Pfu-Polymerase-Puffer, 2,5 U Pfu-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 μL H₂O

Temperaturprofil:

10

25

30

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 30s, 4. 72 °C 2 min, 5. 95 °C 30s, 6. 72 °C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

Anschließend erfolgte eine Reinigung des erhaltenen Fragmentes (~ 1,7 kbp), eine Spaltung mit Hindlll, eine weitere Reinigung des 364-bp-Hindlll-Fragments-carB, gefolgt von einer Spaltung von pBinAHyg mit Hindlll, eine Ligation von 364-bp-HindIII-Fragments-carB in pBinAHyg, eine Transformation des Vektors in Escherichia coli und eine Isolierung des Konstruktes und Bezeichnung als pBinAHygAcarB wie oben beschrieben. Alternativ erfolgte eine partielle Spaltung mit Hindlll und die Klonierung eines größe-15 ren Hindill-Fragmentes aus carB in pBinAHyg zur Herstellung von pBinA-Hyg∆carB.

Erzeugung von carB--Mutanten von Blakeslea trispora 20

Zunächst wurde das Plasmid pBinAHyg∆carB in den Agrobakterienstamm LBA 4404 übertragen, z. B. durch Elektroporation (vgl. oben). Anschließend wurde das Plasmid von Agrobacterium tumefaciens LBA 4404 in Blakeslea trispora ATCC 14272 und in Blakeslea trispora ATCC 14271 übertragen (vgl. oben). Der erfolgreiche Nachweis des Gentransfers in Blakesieslea trispora erfolgte über Polymerase-Kettenreaktion nach folgendem Protokoll:

Ca. 0,5 ug DNA aus Blakeslea trispora ATCC 14272 carB- bzw. ATCC carB- wurden mit 0,25 μM Primer hph forward 14271 CGATGTAGGAGGCGTGGATA-3' (SEQ ID NO 5), 0,25 µM Primer hph reverse 5'-GCTTCTGCGGGCGATTTGTGT-3' (SEQ ID NO 6), 100 μM dNTP, 10 μ L Herculase-Polymerase-Puffer, 2,5 U Herculase-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 μ l H₂O umgesetzt. Temperaturprofil:

1. 95°C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6.72°C 10 min.

Zyklen: (1.-2.) 1x, (3-5.) 30x, (6.) 1x

10

15

Als Negativkontrolle wurde eine Amplifikation des Kanamycinresistenzgens aus Agrobacterium versucht. Dazu wurden folgende PCR-Bedingungen verwendet:

Ca. 0,5 μ g DNA aus Blakesiea trispora ATCC 14272 carB⁻ bzw. ATCC 14271 carB⁻ wurden mit 0,25 μ M Primer nptlll forward 5'-TGAGAATATCACCGGAATTG-3' (SEQ ID NO 7), 0,25 μ M Primer nptlll reverse 5'-AGCTCGACATACTGTTCTTCC-3' (SEQ ID NO 8), 100 μ M dNTP, 10 μ L Herculase-Polymerase-Puffer, 2,5 U Herculase-DNA-Polymerase (Zugabe bei 85 °C, "hot start") und ad 100 μ L H₂O umgesetzt. Temperaturprofil:

1. 95 °C 10 min, 2. 85 °C 5 min, 3. 58 °C 1 min, 4. 72 °C 1 min, 5. 94 °C 1 min, 6. 72 °C 10 min-

20 Zyklen: (1-2.) 1x, (3-5.) 30x, (6.) 1x

C) Produktion von Carotinoiden und Carotinoidvorstufen mit Blakeslea trispora

Zur Produktion der Carotinoide Zeaxanthin, Canthaxanthin, Astaxanthin und Phytoen wurden die entsprechenden gentechnisch veränderten Blakeslea trispora (+) und (-) Stämme fermentiert, das produzierte Carotinoid mittels HPLC Analyse nachgewiesen und isoliert.

Das Flüssigmedium zur Produktion von Carotinoiden enthielt pro Liter: 19 g Maismehl, 44 g Sojamehl, 0,55 g KH₂PO₄, 0,002 g Thiaminhydochlorid, 10 % Sonnenblumenöl. Der pH wurde mit KOH auf 7,5 eingestellt.

Zur Herstellung der Carotinoiden wurden Schüttelkolben mit Sporensuspensionen von (+) und (-) Stämmen der GVO von Blakeslea trispora beimpft. Die Schüttelkolben wurden bei 26 °C mit 250 rpm für 7 Tage inkubiert. Alternativ wurde zu Mischungen der Stämme nach 4 Tagen Trisporsäuren zugegeben und weitere 3 Tage inkubiert. Die Endkonzentration der Trisporsäuren betrug 300 - 400 μg/ml.

Extraktion und Analytik

Extraktion:

- 1. Entnahme von 10 ml Kultursuspension
- 15 2. Zentrifugation, 10 min, 5.000 x g
 - 3. Verwerfen des Überstandes
 - 4. Resuspendierung des Pellets in 1 ml Tetrahydrofuran (THF) durch Vortexen
 - 5. Zentrifugation, 5 min, 5.000 x g
- 20 6. Abnahme der THF-Phase
 - 7. Wiederholung der Schritte 4.-6. (2 x)
 - 8. Vereinigung der THF-Phasen
 - 9. Zentrifugation der vereinigten THF-Phasen 5 min bei 20.000 x g, um Reste der wäßrigen Phase abzutrennen

25

Analytik

Messung von Phytoen mittels HPLC

Säule:

ZORBAX Eclipse XDB-C8, 5 um, 150*4,6 mm

30 Temperatur:

40 °C

Flußrate:

0,5 ml/min

Injektionsvolumen:10 µl

Detektion:

UV 220 nm

Stoppzeit:

12 min

Nachlaufzeit:

0 min

5 Maximaldruck:

350 bar

Eluent A:

50 mM NaH₂PO₄, pH 2,5 mit Perchlorsäure

Eluent B:

Acetonitril

Gradient:

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
10	0	50	50	0,5
	12	50	50	0,5

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe wird durch ein 0,22 µm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 50 - 1000 mg/l eingewogen und in THF gelöst. Als Standard wurde Phytoen verwendet, welches unter den gegebenen Bedingungen eine Retentionszeit von 7,7 min. aufweist.

Messung von Lycopin, β-Carotin, Echinenon, Canthaxanthin, Cryptoxanthin, Zeaxanthin und Astaxanthin mittels HPLC

Säule:

15

Nucleosil 100-7 C18, 250*4,0 mm (Macherey & Nagel)

Temperatur:

25 °C

Flußrate:

1,3 ml/min

25 Injektionsvolumen:10 μl

Detektion:

450 nm

Stoppzeit:

15min

Nachlaufzeit:

2 min

Maximaldruck:

250 bar

30 Eluent A:

10% Aceton, 90% H₂O

Eluent B:

Aceton

\sim					
G	ro	~		n	t٠
u	ıa	u	U	11	١.

15

20

25

. 30

	Zeit [min]	A [%]	B [%]	Fluß [ml/min]
	0	30	70	1,3
	10	5	95	1,3
5	. 12	5 .	95	1,3
	13	30	70	1,3

Als Matrix wurden Extrakte der Fermentationsbrühen verwendet. Vor der HPLC wurde jede Probe wird durch ein 0,22 μm Filter filtriert. Die Proben wurden kühl gehalten und vor Licht geschützt. Zur Kalibrierung wurden jeweils 10 mg eingewogen und in 100 ml THF gelöst. Als Standard wurden folgende Carotinoide mit folgenden Retentionszeiten eingesetzt β-Carotin (12,5 min), Lycopin (11,7 min), Echinenon (10,9 min), Cryptoxanthin (10,5 min), Canthaxanthin (8,7 min), Zeaxanthin (7,6 min) und Astaxanthin (6,4 min) [s. Fig. 23].

Produktion von Zeaxanthin mit gentechnisch veränderten Stämmen von Blakeslea trispora

Nachfolgend wird beispielhaft die Herstellung von Zeaxanthin mit gentechnisch veränderten Organismen (GVO) von Blakeslea trispora beschrieben.

Durch Agrobakterium-vermittelte Transformation wurde der Vektor pBinA-HygBTpTEF1-HPcrtZ in Blakeslea trispora übertragen (s.o.). Ein Hygromycin-resistenter Klon wurde isoliert und auf eine Kartoffel-Glucose-Agarplatte (Merck KGaA, Darmstadt) übertragen.

Nach drei Tagen Inkubation bei 26°C wurde ausgehend von dieser Platte ein Sporensuspension hergestellt. Ein 250-ml-Erlenmeyerkolben ohne Schikanen mit 50 ml Growth-Medium (Maismehl 47 g/l, Sojamehl 23 g/l, KH₂PO₄ 0,5 g/l, Thiamin-HCl 2.0 mg/l, pH mit NaOH vor der Sterilisation auf 6,2–6,7 eingestellt) wurde mit 1x10⁵ Sporen beimpft. Diese Vorkultur inkubierte 48 Stunden bei 26 °C und 250 upm. Für die Hauptkultur

15

20

30

wurde ein 250-ml-Erlenmeyerkolben ohne Schikane enthaltend 40 ml Produktionsmedium mit 4 ml der Vorkultur beimpft und 8 Tage bei 26 °C und 150 upm inkubiert. Das Produktionsmedium enthielt Glucose 50 g/l, Casein Acid Hydrolisate 2 g/l, Hefeextrakt 1 g/l, L-Asparagin 2 g/l, KH₂PO₄ 1,5 g/l, MgSO₄ x 7 H₂O 0,5 g/l, Thiamin-HCl 5 mg/l, Span20 10 g/l, Tween 80 1 g/l, Linolsäure 20 g/l, Maisquellwasser 80 g/l. Nach 72 Stunden erfolgte die Zugabe von Kerosin in einer Endkonzentration von 40 g/l Kerosin.

Nach der Ernte der Kulturen werden die verbliebenen ungefähr 35 ml Kultur mit Wasser auf 40 ml aufgefüllt. Anschließend werden die Zellen im Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin, 3 x bei 1500 bar aufgeschlossen.

Die Suspension mit den aufgeschlossenen Zellen wurde mit 35 ml THF versetzt und 60 min bei RT im Dunkeln bei 250 upm geschüttelt. Danach wurden 2 g NaCl zugegeben und das Gemisch nochmals geschüttelt. Der Extraktionsansatz wurde dann 10 min bei 5000 x g zentrifugiert. Die gefärbte THF-Phase wurde abgenommen, die Zellmasse war vollständig entfärbt.

Die THF-Phase wurde am Rotationsverdampfer bei 30 mbar und 30 °C auf 1 ml eingeengt und danach nochmals in 1 ml THF aufgenommen. Nach Zentrifugation 5 min bei 20 000 x g wurde ein Aliquot der oberen Phase entnommen und durch HPLC analysiert (Fig. 24, Fig. 23).

25 <u>D) Aufarbeitung und Isolierung der Carotinoide bzw. des Nahrungs-</u> mittels

Die oben unter A) angegebenen Kulturbrühen wurden wie nachfolgend aufgearbeitet, um hochreine Carotinoide und ein entsprechendes Nahrungsmittel zu erhalten.

Der Carotinoidgehalt der Kulturbrühen 1, 2, 3 betrug zwischen 0,5 und 1,5 g/L.

D1) Beispiel gemäß Varainte a) IIA und Variante b) IIA bzw. IIB

Die Kulturen mit identischen Medien (insgesamt ca. 1 L) wurden am Ende des Kultivierungszeitraums vereinigt und mit Hilfe eines Dispergiergeräts (Ultra.Turrax ®) homogenisiert.

Die Feststoffkonzentration in den Medien 1 und 2 betrug 37 g/l bzw. 11 g/L. Die Entwässerung der Kulturbrühe erfolgte durch eine Zentrifuge. Bei hohen Zellkonzentrationen bzw. hohem Feststoffgehalt des Mediums kann dieKulturbrühe auch ohne vorherige fest-flüssig-Trennung weiterverarbeitet werden (Medium 3: 127 g Feststoff/L. Nach vorheriger Homogenisation mit einem Dispergiergerät (Ultra-Turrax ®) und unter ständigem Rühren der Suspension wurde die Zellmasse über eine Schlauchpumpe auf den Trockner aufgegeben. Die Eindüsung in den Zylinder des Laborsprühtrockners erfolgte dabei über eine Zweistoffdüse mit dem Durchmesser 2,0 mm. Eingedüst wurde mit 2 bar und 4,5 Nm³/h Stickstoff. Die Temperatur am Eintritt betrug ca. 125°C bis 127°C. Das Trocknungsgas war Stickstoff mit einer Flussrate von 22 Nm³/h. Die Austrittstemperatur betrug zwischen 59°C und 61°C. Bei jeder der drei Fermentationsbrühen konnte am Zyklon des Sprühtrockners rieselfähiges Produkt abgeschieden werden. Die Wandbeläge im Turm (sofern vorhanden) platzten automatisch von der Gefäßwand ab und werden als unproblematisch eingestuft.

25

10

15

20

Es wurden zwischen 8 und 100 g pulvriges Nahrungsmittel erhalten, welches direkt als Tierfuttermittel verwendet werden könnte. Es enthielt ca. 1-10 % Carotinoide bezogen auf das Trockengewicht. Die Restfeuchte betrug weniger als 5%.

10

20

25

30

Beispiel gemäß Variante b) IIC

D2) Extraktion mit Tetrahydrofuran

Die Zellen aus je 40 ml der Kulturbrühen 1, 2, 3 wurden 3 x bei 1500 bar durch einen Hochdruckhomogenisator, Typ Micron Lab 40, Fa. APV Gaulin aufgeschlossen. Je 20 ml der Suspensionen mit den aufgeschlossenen Zellen wurden mit 20 ml Tetrahydrofuran versetzt und 30 min, bei 30°C im Rundschüttler bei 200 Upm geschüttelt. Danach wurden 2 g NaCl zugesetzt und zur Phasentrennung 5 min bei 5000 x g zentrifugiert. Die THF-Phase wurde abgenommen. Danach wurde die wässrige Phase nochmals mit 20 ml THF extrahiert. Die Extrakte wurde vereinigt. Die Carotinoidkonzentration wurde durch HPLC quantifiziert.

D3) Extraktion mit Methylenchlorid

Die Biomassenabtrennung aus der Kulturbrühe (200 mL) erfolgte durch Zentrifugation bei 5.000 x g für 10 min. in einer Laborzentrifuge.

Die abgetrennte Biofeuchtmasse (jeweils ca. 10 g bis 100 g) wurde mit 10 - 100 mL Wasser vermischt, um wasserlösliche Komponenten zu entfernen. Die Biomasse wurde abgetrennt (Laborzentrifuge) und danach mit Dampf (T = 121, t = 30 min, 1 bar) im Autoklaven sterilisiert und so die Zellen aufgeschlossen.

Zu den Zelltrümmern wurden 25 - 250 g Methylenchlorid zugegeben und das Carotinoid aus der Biomasse mittels Ausschütteln extrahiert. Die Biomasse wurde in einer Laborzentrifuge abgetrennt.

Es wurde ein Lösungsmitteltausch von Methylenchlorid zu Methanol durchgeführt, wozu die Carotinoidlösung ca. vier Stunden bei 40°C bis 60°C gehalten und über diesen Zeitraum kontinuierlich mit insgesamt 20 - 200 mL Methanol versetzt wurde. Methylenchlorid wurde dabei als Lö-

30

sungsmittel zurückgewonnen. Erste Carotinoid Kristalle fielen aus. Anschließend wurde langsam, über 6 h auf ca. 10 °C abgekühlt, wobei die Carotinoid Kristalle an Größe und Anzahl zunahmen. Danach wurde die Mutterlauge abfiltriert und die Carotinoid Kristalle getrocknet. Ein Teil der Mutterlauge kann zum Lösungsmitteltausch wiederverwendet werden. Der andere Teil wird destilliert und das so gereinigte Methanol im Lösungsmitteltausch wiederverwendet.

Es wurden 0,0,08 g bis 0,24 g Carotinoid Kristalle erhalten, welche eine Reinheit (HPLC, vgl. oben) von 95 % aufwiesen. Die Ausbeute an Carotinoid Kristallen betrug 80 % bezogen auf die Konzentration an Carotinoid in der Biomasse.

Die abgetrennte methylenchloridfeuchte Biomasse wurde nach Wasserdampfdestillation sprühgetrocknet (T_E = 125 °C, T_A = 60 °C) und kann als Tierfuttermitteladditiv eingesetzt werden.

Hierzu wurde nach vorheriger Homogenisation mit einem Dispergiergerät
(Ultra-Turrax) und unter ständigem Rühren der Suspension die Zellmasse
über eine Schlauchpumpe auf den Trockner aufgegeben.

Die Eindüsung in den Zylinder des Laborsprühtrockners erfolgte dabei über eine Zweistoffdüse mit dem Durchmesser 2,0 mm. Eingedüst wurde mit 2 bar und 4,5 Nm³/h Stickstoff. Die Temperatur am Eintritt betrug ca. 125°C bis 127°C. Das Trocknungsgas war Stickstoff mit einer Flussrate von 22 Nm³/h. Die Austrittstemperatur betrug zwischen 59°C und 61°C. Bei jeder der drei Fermentationsbrühen konnte am Zyklon des Sprühtrockners rieselfähiges Produkt abgeschieden werden. Die Wandbeläge im Turm (sofern vorhanden) platzten automatisch von der Gefäßwand ab und wurden als unproblematisch eingestuft.

Es wurden ca. 2,5 – 25 g pulvriges Nahrungsmittel erhalten, welches direkt als Tierfuttermittel verwendet werden könnte. Es enthielt ca. 0,5% - 1,5% Carotinoide bezogen auf das Trockengewicht. Die Restfeuchte betrug weniger als 5%.

Insgesamt (einschließlich des aufgereinigeten Carotinoid-Nahrungsmittels betrug die Ausbeute an Carotinoid ca, 95 % bezogen auf die Ausgangsmenge Carotinoid in der Kulturbrühe.

WO 2004/063359

10

Patentansprüche

- 1. Verfahren zur Herstellung von Carotinoiden oder deren Vorstufen mittels gentechnisch veränderter Organismen der Gattung Blakeslea umfassend
 - (i) Transformation mindestens einer der Zellen,
- (ii) ggf. Homokaryotisierung der aus (i) erhaltenen Zellen, so dass Zellen entstehen, in denen die Kerne in einem oder in mehreren genetischen Merkmalen alle gleichartig verändert sind und diese genetische Veränderung zur Ausprägung bringen, und
 - (vi) Selektion und Vermehrung der gentechnisch veränderten Zelle oder Zellen,
 - (vii) Kultivierung der gentechnisch veränderten Zellen,
- (viii) Bereitstellung des von den gentechnisch veränderten Zellen produzierten Carotinoids oder der von den gentechnischen veränderten Zellen produzierten Carotinoidvorstufe.
 - 2. Verfahren nach Anspruch 1, dadurch gekennzeichnet, dass es sich um Zellen von Pilzen der Art Blakeslea trispora handelt.
- Verfahren nach Anspruch 1 oder 2, dadurch gekennzeichnet, dass in
 der Transformation (i) ein Vector oder freie Nukleinsäuren verwendet werden.
 - 4. Verfahren nach Anspruch 3, dadurch gekennzeichnet, dass der in der Transformation (i) eingesetzte Vector in das Genom mindestens einer der Zellen integriert wird.

- 5. Verfahren nach Anspruch 4, dadurch gekennzeichnet, dass der in der Transformation (i) eingesetzte Vector einen Promotor und/oder einen Terminator enthält.
- 6. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 5, dadurch gekennzeichnet, dass in der Transformation (i) ein Vector enthaltend den gpd, pcarB, pcarRA und/oder ptef1 Promotor und/oder den trpC Terminator eingesetzt wird.
- 7. Verfahren nach einem der vorhergehenden Ansprüche 3 bis 6, dadurch gekennzeichnet, dass in der Transformation (i) ein Vector enthaltend ein Resistenzgen eingesetzt wird.
 - 8. Verfahren nach Anspruch 7, dadurch gekennzeichnet, dass der in der Transformation (i) eingesetzte Vector ein Hygromycin-Resistenzgen (hph), insbesondere aus E. coli enthält.
- Verfahren nach einem der vorhergehenden Ansprüche 5 8, dadurch gekennzeichnet, dass der gpd Promotor die Sequenz SEQ ID NO: 1 aufweist.
 - 10. Verfahren nach einem der vorhergehenden Ansprüche 5 8, dadurch gekennzeichnet, dass der trpC Terminator die Sequenz SEQ ID NO: 2 aufweist.
 - 11. Verfahren nach einem der vorhergehenden Ansprüche 5 8, dadurch gekennzeichnet, dass der tef1 Promotor die Sequenz SEQ ID NO: 35 aufweist.
- 12. Verfahren nach einem der Ansprüche 6 bis 11, dadurch gekenn zeichnet, dass der gpd Promotor und der trpC Terminator aus Aspergillus nidulans stammen.

- 13. Verfahren nach einem Ansprüche 3 bis 12, dadurch gekennzeichnet, dass der Vector die SEQ ID NO: 3 umfasst.
- 14. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Transformation (i) mittels Agrobakterien, Konjugation, Chemikalien, Elektroporation, Beschuss mit DNA-beladenen Partikeln, Protoplasten oder Mikroinjektion durchgeführt wird.
- 15. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass in der Homokaryontisierung (ii) ein mutagenes Agens eingesetzt wird.
- 16. Verfahren nach Anspruch 15, dadurch gekennzeichnet, dass als mutagenes Agens N-Methyl-N'-nitro-nitrosoguanidin (MNNG), UV-Strahlung oder Röntgenstrahlung eingesetzt wird.
 - 17. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Selektion durch Markierung und/oder Auswahl der einkernigen Zellen erfolgt.
 - 18. Verfahren nach einem der vorhergehenden Ansprüche 1 17, dadurch gekennzeichnet, dass in der Selektion 5-Carbon-5-deazariboflavin (darf) und Hygromycin (hyg) oder 5-Fluororotat (FOA) und Uracil und Hygromycin eingesetzt werden.
- 19. Verfahren nach einem der Ansprüche 3 bis 18, dadurch gekennzeichnet, dass der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinoiden oder deren Vorstufen enthält.
- 20. Verfahren nach einem der Ansprüche 3 bis 19, dadurch gekennzeichnet, dass der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Carotinen oder Xanthophyllen enthält.

21. Verfahren nach einem der Ansprüche 3 bis 20, **dadurch gekennzeichnet**, **dass** der in der Transformation (i) eingesetzte Vector genetische Informationen zur Herstellung von Astaxanthin, Zeaxanthin, Echinenon, β-Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β-Carotin, α-Carotin, Lutein, Phytofluen, Bixin oder Phytoen enthält.

10

15

20

25

- 22. Verfahren zur Bereitstellung mindestens eines hochreinen Carotinoids und eines Nahrungsmittels, enthaltend Carotinoide-produzierende Organismen und mindestens das eine Carotinoid, umfassend nach der Kultivierung von Carotinoide-produzierenden Organismen der Gattung Blakeslea die Schritte
 - I) Abtrennung der Biomasse,
 - IA) ggf. Waschen der Biomasse mit einem Carotinoide nicht lösenden Lösungsmittel, insbesondere Wasser,
 - IB) Sterilisation und Zellaufschluß der Biomasse,
 - IC) ggf. Trocknung und/oder homogene Verteilung und
 - II) partielle Extraktion der Carotinoide aus der aufgeschlossenen Biomasse mittels eines Carotinoide lösenden Lösungsmittels und Trennung des Lösungsmittels von der Biomasse,
 IIA)

10

15

25

- 1) Entfernung von Lösemittelresten aus der Carotinoidhaltigen Biomasse,
- ggf. homogene Suspension der Biomasse mit einem Biomasse-Feststoffgehalt > 10
- 3) Trocknung der Biomasse bzw. Suspension zur Herstellung des Nahrungsmittels,

IIB)

- 1) Kristallisation der Carotinoide aus dem verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- 23. Verfahren nach Anspruch 22, dadurch gekennzeichnet, dass das mindestens eine Carotinoid aus der Gruppe bestehend aus Carotinen und Xanthophyllen ausgewählt ist.
- 24. Verfahren nach Anspruch 22 oder 23, **dadurch gekennzeichnet**, **dass** das mindestens eine Carotinoid aus der Gruppe bestehend aus Asta-xanthin, Zeaxanthin, Echinenon, β-Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon, Lycopin, β-Carotin, Lutein, Phytofluen, Bixin und Phytoen ausgewählt ist.
- 25. Verfahren nach einem der Ansprüche 22 bis 24, dadurch gekennzeichnet, dass das mindestens eine Carotinoid Astaxanthin, Zeaxanthin, Bixin oder Phytoen ist.
 - 26. Verfahren nach einem der Ansprüche 22-25, dadurch gekennzeichnet, dass die Sterilisation und der Zellaufschluß mittels Wasserdampf oder Mikrowellenstrahlung durchgeführt werden.
 - 27. Verfahren nach einem der Ansprüche 22-26, dadurch gekennzeichnet, dass die Extraktion der Carotinoide aus der Biomasse mittels Me-

- thylenchlorid oder überkritischem Kohlendioxid oder Tetrahydrofuran durchgeführt wird.
- 28. Verfahren nach Anspruch 27, dadurch gekennzeichnet, dass die im überkritischen Kohlendioxid gelösten Carotinoide direkt isoliert werden oder in Methylenchlorid aufgenommen werden.
- 29. Verfahren nach einem der Ansprüche 22-28, dadurch gekennzeichnet, dass die Extraktion der Carotinoide aus der Biomasse ein oder ggf. mehrstufig erfolgt.
- 30. Verfahren nach einem der Ansprüche 22-29, dadurch gekennzeichnet, dass die Entfernung von Lösungsmitteln aus der Biomasse im
 Schritt IA1) mittels Wasserdampf-Destillation
 - 31. Verfahren nach einem der Ansprüche 22-30, dadurch gekennzeichnet, dass die Trocknung in Schritt IIA3) mittels Sprühtrocknung oder Kontakttrocknung durchgeführt wird.
- 32. Verfahren nach einem der vorhergehenden Ansprüche, dadurch gekennzeichnet, dass die Kristallisation im Schritt IIB1) durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel erfolgt.
- 33. Verfahren nach Anspruch 32, dadurch gekennzeichnet, dass der Austausch des verwendeten Lösungsmittels gegen Wasser oder einen niederen Alkohol, insbesondere Methanol erfolgt.
 - 34. Verfahren nach Anspruch 13, dadurch gekennzeichnet, dass der gentechnisch veränderte Organismus der Gattung Blakeslea durch Transformation mit einem Vector, der eine Sequenz aus der Gruppe bestehend aus den SEQ ID NO: 37 51 und 62 aufweist, herstellbar ist.

- 35. Verfahren zur Herstellung eines Nahrungsmittels enthaltend Organismen der Gattung Blakeslea und mindestens ein Carotinoid, umfassend nach der Kultivierung von Carotinoide-produzierenden Organismen der Gattung Blakeslea die Schritte
- I) Homogene Suspendierung der Feststoffe der Kulturbrühe und
 - IIA) bei einem Biomasse-Feststoffgehalt der Kulturbrühe von > 2 %
 - 1) ggf. Konzentration der Kulturbrühe auf einen Feststoffgehalt < 50 % und
 - Trocknung der Kulturbrühe zur Herstellung des Nahrungsmittels

oder

15

10

- IIB) bei einem Feststoffgehalt von < 2 % der Kulturbrühe,
 - Konzentration der Kulturbrühe auf einen Feststoffgehalt > 2 % und < 50 % und
 - Trocknung der Suspension zur Herstellung des Nahrungsmittels,

oder

- IIC) unabhängig vom Feststoffgehalt der Kulturbrühe,
 - 1) Abtrennung der Biomasse,

25

30

- ggf. Waschen der Biomasse mit Carotinoide nicht lösenden Lösungsmitteln, insbesondere Wasser,
- 3) Sterilisation und Zellaufschluß,
- 4) ggf. Trocknung und homogene Verteilung,
- partielle Extraktion der Carotinoide aus der Biomasse mittels eines Carotinoide lösendes Lösungsmittels,

15

- 5a) Abtrennung der Carotinoid-haltigen Biomasse vom Carotinoid-haltigen Lösungsmittel,
- 5b) Entfernung von Lösemittelresten aus der Biomasse und
- 5c) Trocknung der Biomasse zur Herstellung des Nahrungsmittels,
- 6) Kristallisation der Carotinoide aus dem in 5a) verwendeten Lösungsmittel und Isolierung der Carotinoid-Kristalle, insbesondere durch Filtration.
- 36. Verfahren nach Anspruch 35, dadurch gekennzeichnet, dass das mindestens eine Carotinoid aus der Gruppe bestehend aus Carotinen und Xanthophyllen ausgewählt ist.
 - 37. Verfahren nach Anspruch 35 oder 36, dadurch gekennzeichnet, dass das mindestens eine Carotinoid aus der Gruppe bestehend aus Astaxanthin, Zeaxanthin, Echinenon, β-Cryptoxanthin, Andonixanthin, Adonirubin, Canthaxanthin, 3-Hydroxyechinenon, 3'-Hydroxyechinenon,
 Lycopin, β-Carotin, Lutein, Bixin, Phytoen ausgewählt ist.
 - 38. Verfahren nach einem der Ansprüche 35-37, dadurch gekennzeichnet, dass das mindestens eine Carotinoid Astaxanthin, Zeaxanthin, *Bi-xin* oder Phytoen ist.
 - 39. Verfahren nach einem der Ansprüche 35-38, dadurch gekennzeichnet, dass die Sterilisation und der Zellaufschluß im Schritt II3) mittels Wasserdampf oder Mikrowellenstrahlung durchgeführt wird.
- 40. Verfahren nach einem der Ansprüche 35-39, dadurch gekennzeichnet, dass die Extraktion der Carotinoide aus der Biomasse im Schritt
 IIC5) mittels Methylenchlorid oder überkritischen Kohlendioxid durchgeführt wird.

- 41. Verfahren nach Anspruch 40, dadurch gekennzeichnet, dass die im überkritischen Kohlendioxid gelösten Carotinoide direkt isoliert werden oder in Methylenchlorid aufgenommen werden..
- 42. Verfahren nach einem der Ansprüche 35-41, dadurch gekennzeichnet, dass die Extraktion der Carotinoide aus der Biomasse ein- oder ggf. mehrstufig erfolgt.
 - 43. Verfahren nach einem der Ansprüche 35-42, dadurch gekennzeichnet, dass die Entfernung von Lösungsmitteln aus der Biomasse im Schritt IIC5b) mittels Wasserdampf-Destillation.
- 44. Verfahren nach einem der Ansprüche 35-43, dadurch gekennzeichnet, dass die Trocknung in einem der Schritte IIA1), IIB2) oder IIC5c) mittels Sprühtrocknung oder Kontakt durchgeführt wird.
 - 45. Verfahren nach einem der Ansprüche 35-44, dadurch gekennzeichnet, dass die Kristallisation im Schritt IIC6) durch graduellen Lösungsmittelaustausch gegen ein Carotinoide nicht lösendes Lösungsmittel erfolgt.

- 46. Verfahren nach Anspruch 45, dadurch gekennzeichnet, dass der Austausch des verwendeten Lösungsmittels gegen Wasser oder einen niederen Alkohol, insbesondere Methanol erfolgt.
- 47. Verfahren nach einem der Ansprüche 35-46, dadurch gekennzeichnet, dass der gentechnisch veränderte Organismus der Gattung Blakeslea durch Transformation mit einem Vector, der eine Sequenz aus der Gruppe bestehend aus den SEQ ID NO: 37 51 und 62 aufweist, herstellbar ist.
- 48. Nahrungsmittel, insbesondere Tierfuttermittel herstellbar nach einem der Verfahren der Ansprüche 1 bis 47.

- 49. Nahrungsergänzungsmittel, insbesondere Tierfutterergänzungsmittel herstellbar nach einem der Verfahren der Ansprüche 1 bis 47.
- 50. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß Nahrungsmittel und Tierfuttermittel aus einer Fermentation erhältlich sind.
- 51. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß Nahrungsergänzungsmittel und Tierfutterergänzungsmittel aus einer Fermentation erhältlich sind.
- 52. Verfahren nach einem der Ansprüche 1-49 dadurch gekennzeichnet, daß mindestens zwei Produkte aus der Gruppe Nahrungsmittel, Nahrungsergänzungsmittel, Tierfuttermittel und Tierfutterergänzungsmittel aus einer Fermentation erhältlich sind.
 - 53. Verwendung der nach einem der Verfahren der Ansprüche 1 bis 14 erhältlichen Carotinoide zur Herstellung von kosmetischen, pharmazeutischen, dermatologischen Zubereitungen, Nahrungsmitteln, Nahrungsergänzungsmitteln, Tierfuttermittel oder Tierfutterergänzungsmittel.

Fig. 1: Vektor pANsCos1

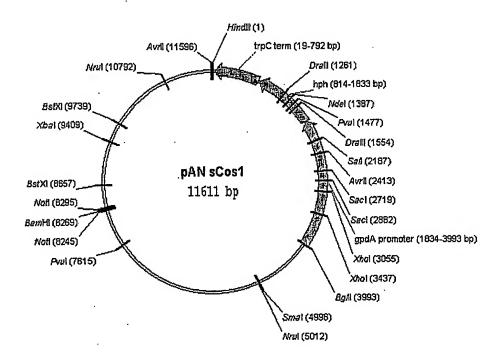


Fig. 2: Vektor pBinAHyg

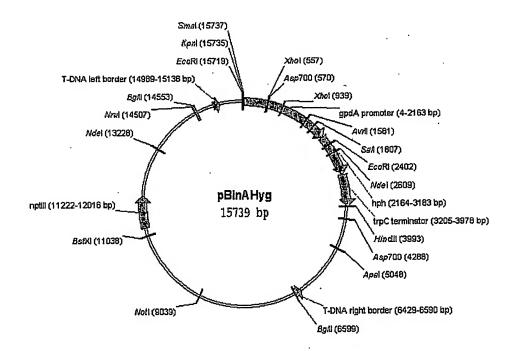


Fig. 3: Gels des Ergebnis einer PCR Spur:

1 2 3 4 5

Fig. 4: Plasmid pBinAHygBTpyrG-SCO

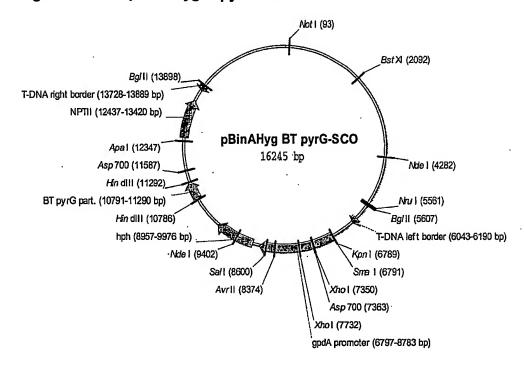


Fig. 5: Plasmid pBinAHygBTpTEF1-HPcrtZ

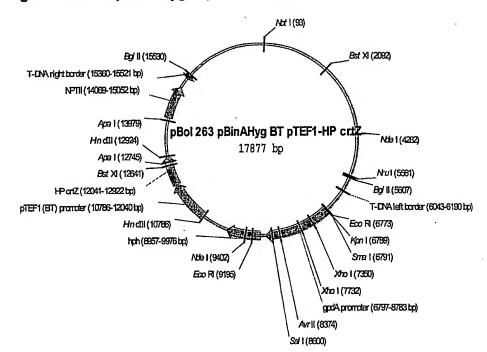


Fig. 6: Plasmid pBinAHyg-BTpcarRA-HPcrtZ

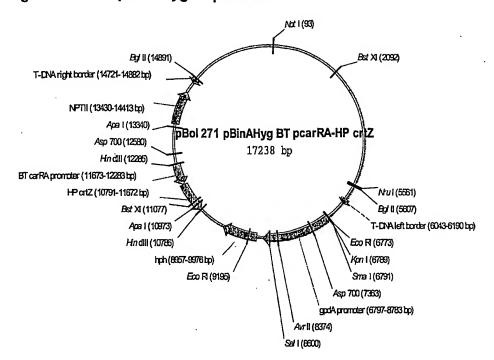


Fig. 7: Plasmid pBinAHygBTpcarB-HPcrtZ

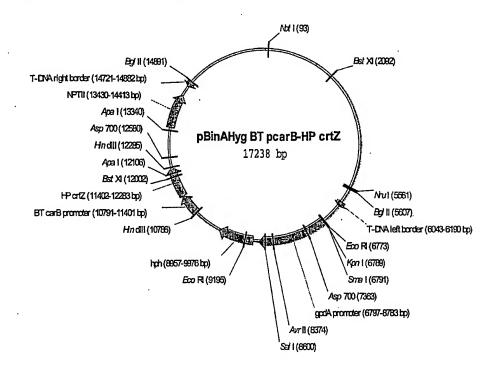


Fig. 8: Plasmid p-carRA-HPcrtZ-TAG-3'carA-IR

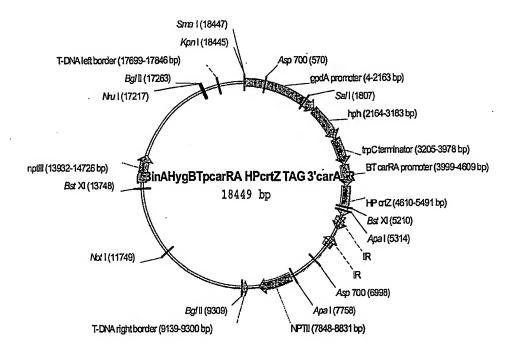


Fig. 9: Plasmid p-carRA-HPcrtZ-GCG-3'carA-IR

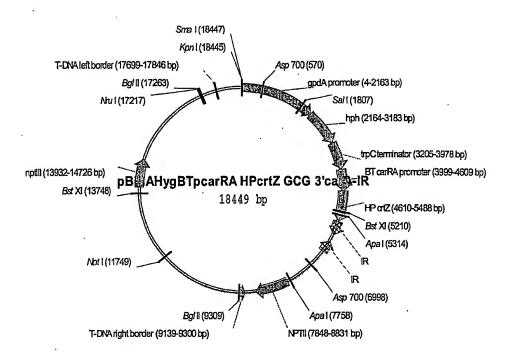


Fig. 10: Plasmid pBinAHygBTpTEF1-EUcrtZ

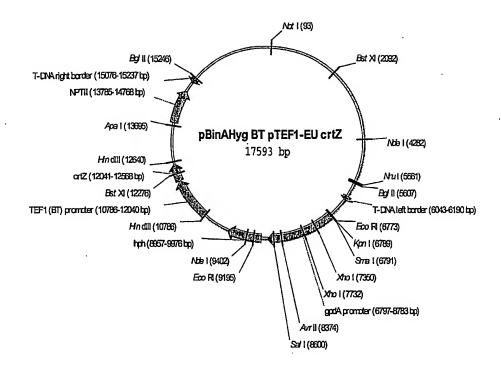


Fig. 11: Plasmid pBinAHygBTpcarRA-EUcrtZ

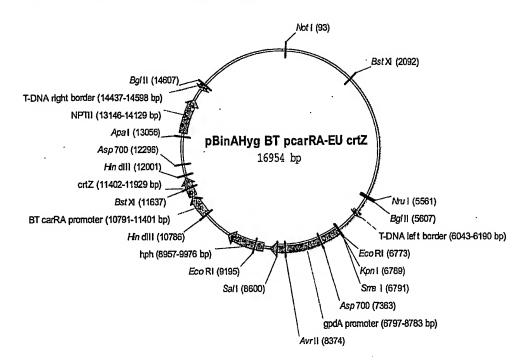


Fig. 12: Plasmid pBinAHygBTpcarB-EUcrtZ

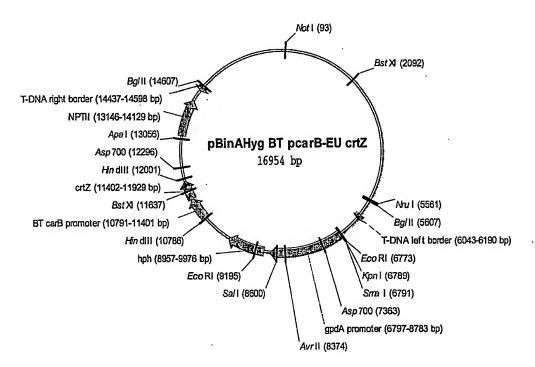


Fig. 13: Plasmid p-BinAHyg-gpdA-HPcrtZ

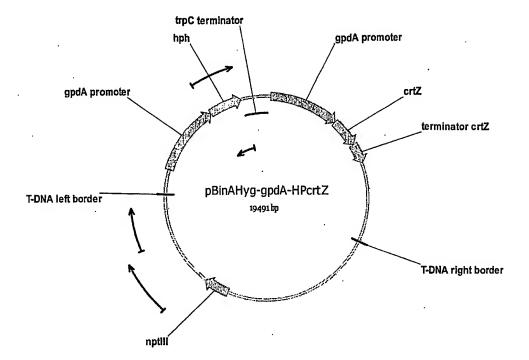


Fig. 14: Plasmid pBinAHyg-carRcrtZcarA

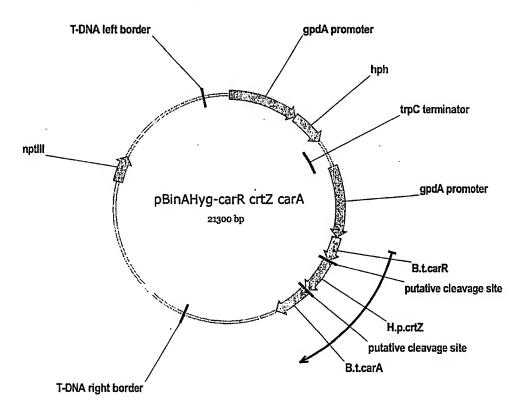


Fig. 15: Plasmid pBinAHyg-BTpTEF1-NPcrtW

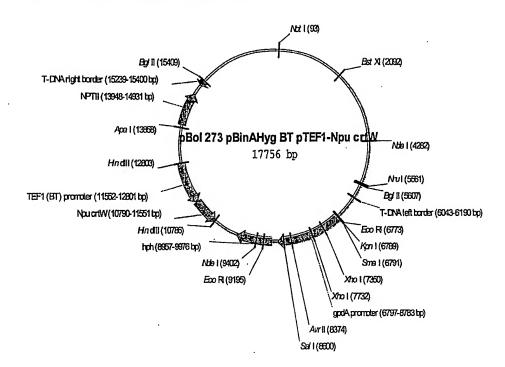


Fig. 16: Plasmid pBinAHyg_BTpcarRA_NPcrtW

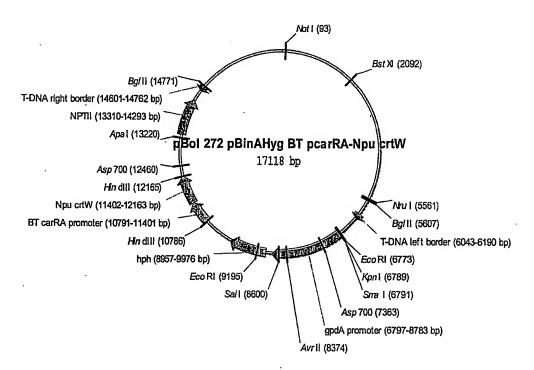


Fig. 17: Plasmid pBinAHyg-BTpcarB-NPcrtW

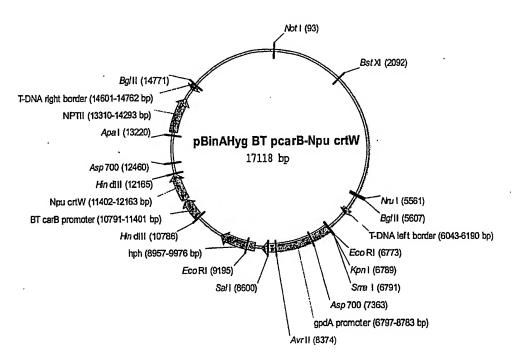
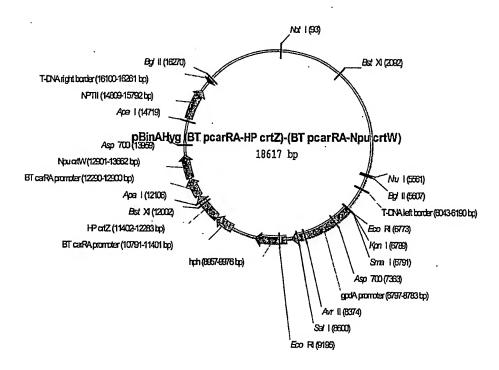


Fig. 18: Plasmid pBinAHygBTpcarRA-HPcrtZ-BTpcarRA-NpucrtW



PCT/EP2004/000099

WO 2004/063359

Fig. 19: Plasmid pBinAHygBTpcarRA-EUcrtZ-BTpcarRA-NpucrtW

19/24

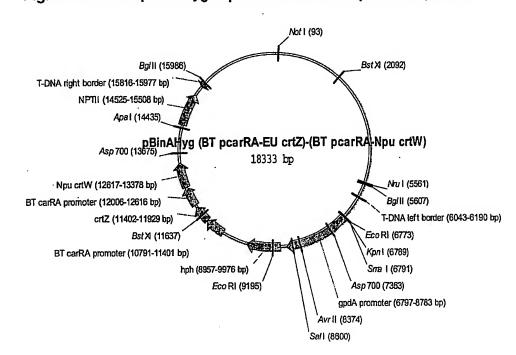


Fig. 20: carB

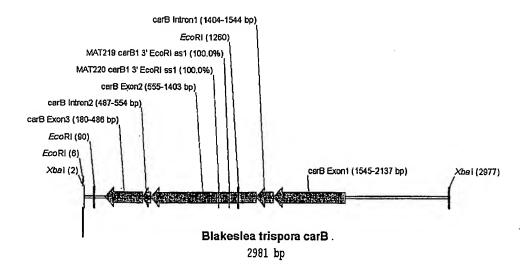


Fig. 21: CDS von carB

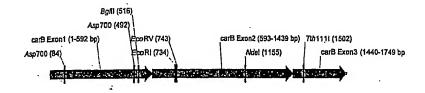


Fig. 22: Vektor pBinAHyg∆carB

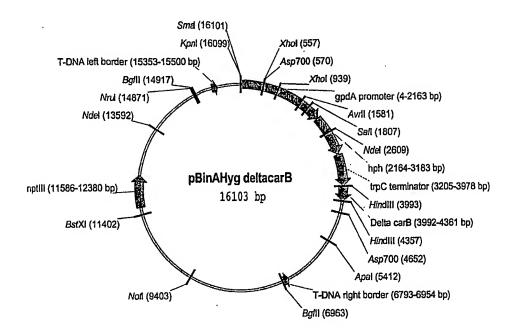


Fig. 23: HPLC Standard

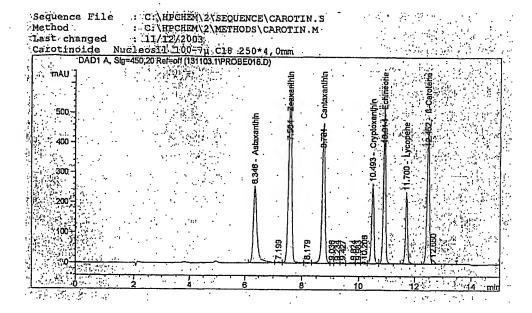
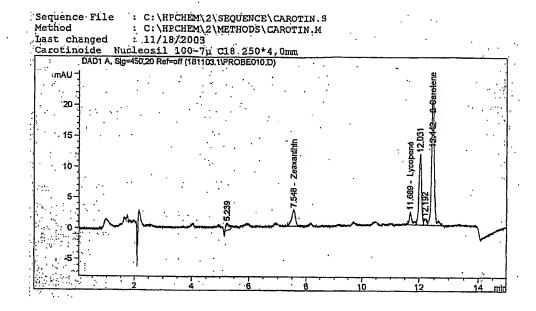


Fig. 24: HPLC



1/357

SEQUENCE LISTING

<110>	BASF AG	
<120>	Verfahren zur Herstellung von Carotinoiden oder deren Vorstufe mittels gentechnisch veränderter Organismen der Gattung Blakeslea, mit dem Verfahren hergestellte Carotinoide oder der Vorstufen und deren Verwendung	•
<130>	BASF/NAE877/03	
<160>	80	
<170>	PatentIn version 3.2	
<210>	1	
<211>	2160	
<212>	DNA	
<213>	Artificial	
<220>		
<223>	Promotor	
<400>	1	
ctttcga	cac tgaaatacgt cgagectget cegettggaa geggegagga geetegteet	60
gtcacaa	cta ccaacatgga gtacgataag ggccagttce gccagetcat taagagccag	120
ttcatgg	gcg ttggcatgat ggccgtcatg catctgtact tcaagtacac caacgctctt	180
ctgatcc	agt cgatcatccg ctgaaggcgc tttcgaatct ggttaagatc cacgtcttcg 2	240
ggaagcc	agc gactggtgac ctccagcgtc cctttaaggc tgccaacagc tttctcagcc	300
agggcca	gcc caagacegac aaggeeteee teeagaaege egagaagaae tggaggggtg	360
gtgtcaa	gga ggagtaagct cettattgaa gteggaggae ggageggtgt caagaggata 🦸	120
ttcttcg	act ctgtattata gataagatga tgaggaattg gaggtagcat agcttcattt 4	80

ggatttgctt tccaggctga gactctagct tggagcatag agggtccttt ggctttcaat 540

WO 2004/063359 PCT/EP2004/000099 2/357

attctcaagt	atctcgagtt	tgaacttatt	ccctgtgaac	cttttattca	ccaatgagca	- 600
ttggaatgaa	catgaatctg	aggactgcaa	tegecatgag	gttttcgaaa	tacatccgga	. 660
tgtcgaaggc	ttggggcacc	tgcgttggtt	gaatttagaa	cgtggcacta	ttgatcatcc	720
gatagctctg	caaagggcgt	tgcacaatgc	aagtcaaacg	ttgctagcag	ttccaggtgg	780
aatgtťatga	tgagcattgt	attaaatcag	gagatatagc	atgatctcta	gttagctcac	840
cacaaaagtc	agacggcgta	accaaaagtc	acacaacaca	agctgtaagg	atttcggcac	900
ggctacggaa	gacggagaag	ccaccttcag	tggactcgag	taccatttaa	ttctatttgt	960
gtttgatcga	gacctaatac	agcccctaca	acgaccatca	aagtcgtata	gctaccagtg	1020
aggaagtgga	ctcaaatcga	cttcagcaac	atctcctgga	taaactttaa	gcctaaacta	1080
tacagaataa	gataggtgga	gagcttatac	cgagctccca	aatctgtcca	gatcatggtt	1140
gaccggtgcc	tggatcttcc	tatagaatca	tccttattcg	ttgacctagc	tgattctgga	1200
gtgacccaga	gggtcatgac	ttgagcctaa	aatccgccgc	ctccaccatt	tgtagaaaaa	1260
tgtgacgaac	tcgtgagctc	tgtacagtga	ccggtgactc	tttctggcat	gcggagagac	1320
ggacggacgc	agagagaagg	gctgagtaat	aagccactgg	ccagacagct	ctggcggctc	1380
tgaggtgcag	tggatgatta	ttaatccggg	accggccgcc	cetecgecee	gaagtggaaa	1440
ggctggtgtg	cccctcgttg	accaagaatc	tattgcatca	tcggagaata	tggagcttca	1500
tcgaatcacc	ggcagtaagc	gaaggagaat	gtgaagccag	gggtgtatag	ccgtcggcga	1560
aatagcatgc	cattaaccta	ggtacagaag	tccaattgct	tccgatctgg	taaaagattc	1620
acgagatagt	accttctccg	aagtaggtag	agcgagtacc	cggcgcgtaa	gctccctaat	1680
tggcccatcc	ggcatctgta	gggcgtccaa	atatcgtgcc	tetectgett	tgcccggtgt	1740
atgaaaccgg	aaaggccgct	caggagctgg	ccagcggcgc	agaccgggaa	cacaagctgg	1800

agetttgece egetgtet etgeactega eetgetgagg teecteagte eetggtagge 1860

agetttgece egtetgteeg eeeggtgtgt eggeggggtt gacaaggteg ttgegteagt 1920

ccaacatttg ttgecatatt tteetgetet eeecaceage tgetettte ttttetettt 1980

ctttteecat etteagtata tteatettee eatecaagaa eetttatte eeetaagtaa 2040

gtaetttget acatecatae teeateette eeatecetta tteetttgaa eettteagtt 2100

egagetttee eactteateg eagettgaet aacagetaee eegettgage agacateace 2160

<210> 2

<211> 774

<212> DNA

<213> Artificial

<220>

<223> Terminator

<220>

<221> misc_feature

<222> (267)..(267)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (475)..(475)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (566)..(566)

<223> n is a, c, g, or t

<400> 2

cgatccactt aacgttactg aaatcatcaa acagcttgac gaatctggat ataagatcgt

tggtgtcgat gtcagctccg gagttgagac aaatggtgtt caggatctcg ataagatacg

120

4/3

ttcatttgtc caagcagcaa agagtgcctt ctagtgattt aatagctcca tgtcaacaag 240 aataaaacgc gttttcgggt ttacctcttc cagatacagc tcatctgcaa tgcattaatg cattgactgc aacctagtaa cgccttncag gctccggcga agagaagaat agcttagcag 300 agctattttc attttcggga gacgagatca agcagatcaa cggtcgtcaa gagacctacg 360 420 agactgagga atccgctctt ggctccacgc gactatatat ttgtctctaa ttgtactttg acatgctcct cttctttact ctgatagctt gactatgaaa attccgtcac cagcncctgg 480 540 gttcgcaaag ataattgcat gtttcttcct tgaactctca agcctacagg acacacattc 600 -atcqtaggta taaacctcga aatcanttcc tactaagatg gtatacaata gtaaccatgc 660 atgqttqcct agtgaatgct ccgtaacacc caatacgccg gccgaaactt ttttacaact ctcctatgag tcgtttaccc agaatgcaca ggtacacttg tttagaggta atccttcttt 720 774 ctagctagaa gtcctcgtgt actgtgtaag cgcccactcc acatctccac tcga

<210> 3

<211> 15739

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

. <222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 3

gatctttcga	cactgaaata	cgtcgagcct	gctccgcttg	gaagcggcga	ggagcctcgt	60
cctgtcacaa	ctaccaacat	ggagtacgat	aagggccagt	teegeeaget	cattaagagc	. 120
cagttcatgg	gcgttggcat	gatggccgtc	atgcatctgt	acttcaagta	caccaacgct	180
cttctgatcc	agtcgatcat	ccgctgaagg	cgctttcgaa	tctggttaag	atccacgtct	240
tegggaagee	agcgactggt	gacctccagc	gtccctttaa	ggctgccaac	agctttctca	. 300
gccagggcca	gcccaagacc	gacaaggcct	ccctccagaa	cgccgagaag	aactggaggg	360
gtggtgtcaa	ggaggagtaa	gctccttatt	gaagtcggag	gacggagcgg	tgtcaagagg	420
atattcttcg	actctgtatt	atagataaga	tgatgaggaa	ttggaggtag	catagcttca	480
tttggatttg	ctttccaggc	tgagactcta	gcttggagca	tagagggtcc	tttggctttc	540
aatattctca	agtatctcga	gtttgaactt	attccctgtg	aaccttttat	tcaccaatga	600
gcattggaat	gaacatgaat	ctgaggactg	caatcgccat	gaggttttcg	aaatacatcc	660
ggatgtcgaa	ggcttggggc	acctgcgttg	gttgaattta	gaacgtggca	ctattgatca	720
tccgatagct	ctgcaaaggg	cgttgcacaa	tgcaagtcaa	acgttgctag	cagttccagg	780
tggaatgtta	tgatgagcat	tgtattaaat	caggagatat	agcatgatct	ctagttagct	840
caccacaaaa	gtcagacggc	gtaaccaaaa	gtcacacaac	acaagctgta	aggatttcgg	900
cacggctacg	gaagacggag	aagccacctt	cagtggactc	gagtaccatt	taattctatt	960
tgtgtttgat	cgagacctaa	tacageceet	acaacgacca	tcaaagtcgt	atagctacca	1020
gtgaggaagt	ggactcaaat	cgacttcagc	aacatctcct	ggataaactt	taagcctaaa	1080

WO 2004/063359 PCT/EP2004/000099 6/357

ctatacagaa	taagataggt	ggagagctta	taccgagctc	ccaaatctgt	ccagatcatg	1140
gttgaccggt	gcctggatct	tcctatagaa	tcatccttat	tcgttgacct	agctgattct	1200
ggagtgaccc	agagggtcat	gacttgagcc	taaaatccgc	cgcctccacc	atttgtagaa	.1260
aaatgtgacg	aactcgtgag	ctctgtacag	tgaccggtga	ctctttctgg	catgcggaga	1320
gacggacgga	cgcagagaga	agggctgagt	aataagccac	tggccagaca	gctctggcgg	1380
ctctgaggtg	cagtggatga	ttattaatcc	gggaccggcc	gcccctccgc	cccgaagtgg	1440
aaaggctggt	gtgcccctcg	ttgaccaaga	atctattgca	tcatcggaga	atatggagct	1500
tcatcgaatc	accggcagta	agcgaaggag	aatgtgaagc	caggggtgta	tagccgtcgg	1560
cgaaatagca	tgccattaac	ctaggtacag	aagtccaatt	gcttccgatc	tggtaaaaga	1620
ttcacgagat	agtaccttct	ccgaagtagg	tagagcgagt	acccggcgcg	taagctccct	1680
aattggccca	tccggcatct	gtagggcgtc	caaatatcgt	gcctctcctg	ctttgcccgg	1740
tgtatgaaac	cggaaaggcc	gctcaggagc	tggccagcgg	cgcagaccgg	gaacacaagc	1800
tggcagtcga	cccatccggt	gctctgcact	cgacctgctg	aggtccctca	gtccctggta	1860
ggcagctttg	cccgtctgt	ccgcccggtg	tgtcggcggg	gttgacaagg	tcgttgcgtc	1920
agtccaacat	ttgttgccat	attttcctgc	tctccccacc	agctgctctt	ttcttttctc	1980
tttctttcc	catcttcagt	atattcatct	tcccatccaa	gaacctttat	ttcccctaag	2040
taagtacttt	gctacatcca	tactccatcc	ttcccatccc	ttattccttt	gaacctttca	2100
gttcgagctt	tcccacttca	tegeagettg	actaacagct	accccgcttg	agcagacatc	2160
accatgcctg	aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	gttcgacagc	2220
gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	cttcgatgta	. 2280
ggagggcgtg	gatatgtcct	gcgggtaaat	agctgcgccg	atggtttcta	caaagatcgt	2340

tatgtttatc	ggcactttgc	atcggccgcg	ctcccgattc	cggaagtgct	tgacattggg	2400
gaattcagcg	agagcctgac	ctattgcatc	tcccgccgtg	cacagggtgt	cacgttgcaa	2460
gacctgcctg	aaaccgaact	gcccgctgtt	ctgcagccgg	tcgcggaggc	catggatgcg	2520
atcgctgcgg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaatc	2580
ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	tgtgtatcac	2640
tggcaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	cgatgagctg	2700
atgctttggg	ccgaggactg	ccccgaagtc	cggcacctcg	tgcacgcgga	tttcggctcc	2760
aacaatgtcc	tgacggacaa	tggccgcata	acageggtea	ttgactggag	cgaggcgatg	2820
ttcggggatt	cccaatacga	ggtcgccaac	atcttcttct	ggaggccgtg	gttggcttgt	2880
atggagcagc	agacgcgcta	cttcgagcgg	aggcatccgg	agcttgcagg	atcgccgcgg	2940
ctccgggcgt	atatgctccg	cattggtctt	gaccaactct	atcagagctt	ggttgacggc	3000
aatttcgatg	atgcagcttg	ggcgcagggt	cgatgcgacg	caatcgtccg	atccggagcc	3060
gggactgtcg	ggcgtacaca	aatcgcccgc	agaagcgcgg	ccgtctggac	cgatggctgt	3120
gtagaagtac	togoogatag	tggaaaccga	cgccccagca	ctcgtccgag	ggcaaaggaa	3180
tagagtagat	gccgaccgcg	ggatcgatcc	acttaacgtt	actgaaatca	tcaaacagct	3240
tgacgaatct	ggatataaga	tcgttggtgt	cgatgtcagc	tccggagttg	agacaaatgg	3300
tgttcaggat	ctcgataaga	tacgttcatt	tgtccaagca	gcaaagagtg	ccttctagtg	3360
atttaatagc	tccatgtcaa	caagaataaa	acgcgttttc	gggtttacct	cttccagata	3420
cageteatet	gcaatgcatt	aatgcattga	ctgcaaccta	gtaacgcctt	ncaggctccg	3480
gcgaagagaa	gaatagctta	gcagagctat	tttcattttc	gggagacgag	atcaagcaga	3540
tcaacggtcg	tcaagagacc	tacgagactg	aggaatccgc	tcttggctcc	acgcgactat	3600

WO 2004/063359 PCT/EP2004/000099 8/357

atatttgtct ctaattgtac tttgacatgc tcctcttctt tactctgata gcttgactat 3660 gaaaattccg tcaccagcnc ctgggttcgc aaagataatt gcatgtttct tccttgaact 3720 ctcaagccta caggacacac attcatcgta ggtataaacc tcgaaatcan ttcctactaa 3780 gatggtatac aatagtaacc atgcatggtt gcctagtgaa tgctccgtaa cacccaatac 3840 geoggeogaa aetttttac aacteteeta tgagtegttt acceagaatg cacaggtaca 3900 cttgtttaga ggtaatcctt ctttctagct agaagtcctc gtgtactgtg taagcgccca 3960 ctccacatct ccactegacc tgcaggcatg caagcttggc gtaatcatgg tcatagctgt 4020 ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc ggaagcataa 4080 agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg ttgcgctcac 4140 tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc ggccaacgcg 4200 cggggagagg cggtttgcgt attgggccaa agacaaaagg gcgacattca accgattgag 4260 ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca ccgtcaccga 4320 cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta ccattagcaa 4380 ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta gcgacagaat 4440 caagtttgcc tttagcgtca gactgtagcg cgttttcatc ggcattttcg gtcatagccc 4500 ccttattagc gtttgccatc ttttcataat caaaatcacc ggaaccagag ccaccaccgg 4560 aaccgcctcc ctcagagccg ccaccctcag aaccgccacc ctcagagcca ccaccctcag 4620 4680 agccgccacc agaaccacca ccagagccgc cgccagcatt gacaggaggc ccgatctagt aacatagatg acaccgcgcg cgataattta tcctagtttg cgcgctatat tttgttttct 4740 atcgcgtatt aaatgtataa ttgcgggact ctaatcataa aaacccatct cataaataac 4800 gtcatgcatt acatgttaat tattacatgc ttaacgtaat tcaacagaaa ttatatgata 4860

WO 2004/063359 PCT/EP2004/000099 9/357

atcategeaa ga	accggcaac	aggattcaat	cttaagaaac	tttattgcca	aatgtttgaa	4920
cgatcgggga to	catccgggt	ctgtggcggg	aactccacga	aaatatccga	acgcagcaag	4980
atatcgcggt go	catctcggt	cttgcctggg	cagtcgccgc	cgacgccgtt	gatgtggacg	5040
ccgggcccga to	catattgtc	gctcaggatc	gtggcgttgt	gcttgtcggc	cgttgctgtc	5100
gtaatgatat co	ggcaccttc	gaccgcctgt	tccgcagaga	tcccgtgggc	gaagaactcc	5160
agcatgagat co	cccgcgctg	gaggatcatc	cagccggcgt	cccggaaaac	gattccgaag	5220
cccaaccttt ca	atagaaggc	ggcggtggaa	tcgaaatctc	gtgatggcag	gttgggcgtc	5280
gcttggtcgg to	catttcgaa	ccccagagtc	ccgctcagaa	gaactcgtca	agaaggcgat	5340
agaaggcgat go	cgctgcgaa	tcgggagcgg	cgataccgta	aagcacgagg	aagcggtcag	5400
cccattcgcc go	ccaagctct	tcagcaatat	cacgggtagc	caacgctatg	tectgatage	5 4 60
ggtccgccac ac	cccagccgg	ccacagtcga	tgaatccaga	aaagcggcca	ttttccacca	5520
tgatattcgg ca	aagcaggca	tegecatggg	tcacgacgag	atcatcgccg	tegggeatge	5580
gcgccttgag co	ctggcgaac	agttcggctg	gcgcgagccc	ctgatgctct	tegtecagat	5640
catcctgatc ga	acaagaccg	gcttccatcc	gagtacgtgc	tcgctcgatg	cgatgtttcg	5700
cttggtggtc ga	aatgggcag	gtagccggat	caagcgtatg	cagccgccgc	attgcatcag	5760
ccatgatgga ta	actttctcg	gcaggagcaa	ggtgagatga	caggagatcc	tgccccggca	5820
cttcgcccaa ta	agcagccag	tcccttcccg	cttcagtgac	aacgtcgagc	acagetgege	5880
aaggaacgcc c	gtcgtggcc	agccacgata	gccgcgctgc	ctcgtcctgc	agttcattca	5940
gggcaccgga ca	aggtcggtc	ttgacaaaaa	gaaccgggcg	cccctgcgct	gacagccgga	6000
acacggcggc a	tcagagcag	ccgattgtct	gttgtgccca	gtcatagccg	aatagcctct	6060
ccacccaagc g	gccggagaa	cctgcgtgca	atccatcttg	ttcaatcatg	cgaaacgatc	6120

WO 2004/063359 PCT/EP2004/000099 10/357

cagatccggt	gcagattatt	tggattgaga	gtgaatatga	gactctaatt	ggataccgag	6180
gggaatttat	ggaacgtcag	tggagcattt	ttgacaagaa	atatttgcta	gctgatagtg	6240
accttaggcg	acttttgaac	gcgcaataat	ggtttctgac	gtatgtgctt	agctcattaa	6300
actccagaaa	cccgcggctg	agtggctcct	tcaacgttgc	ggttctgtca	gttccaaacg	6360
taaaacggct	tgtcccgcgt	catcggcggg	ggtcataacg	tgactccctt	aattctccgc	6420
tcatgatcag	attgtcgttt	cccgccttca	gtttaaacta	tcagtgtttg	acaggatata	6480
ttggcgggta	aacctaagag	aaaagagcgt	ttattagaat	aatcggatat	ttaaaagggc	6540
gtgaaaaggt	ttatccgttc	gtccatttgt	atgtgcatgc	caaccacagg	gttccccaga	6600
tetggegeeg	gccagcgaga	cgagcaagat	tggccgccgc	ccgaaacgat	ccgacagcgc	6660
gcccagcaca	ggtgcgcagg	caaattgcac	caacgcatac	agegeeagea	gaatgccata	6720
gtgggcggtg	acgtcgttcg	agtgaaccag	atcgcgcagg	aggcccggca	gcaccggcat	6780
aatcaggccg	atgccgacag	cgtcgagcgc	gacagtgctc	agaattacga	tcaggggtat	6840
gttgggtttc	acgtctggcc	tccggaccag	cctccgctgg	tccgattgaa	cgcgcggatt	6900
ctttatcact	gataagttgg	tggacatatt	atgtttatca	gtgataaagt	gtcaagcatg	6960
acaaagttgc	agccgaatac	agtgatccgt	gccgccctgg	acctgttgaa	cgaggtcggc	7020
gtagacggtc	tgacgacacg	caaactggcg	gaacggttgg	gggttcagca	gccggcgctt	7080
tactggcact	tcaggaacaa	gcgggcgctg	ctcgacgcac	tggccgaagc	catgctggcg	7140
gagaatcata	cgcattcggt	gccgagagcc	gacgacgact	ggcgctcatt	tctgatcggg	7200
aatgcccgca	gcttcaggca	ggcgctgctc	gcctaccgcg	atggcgcgcg	catccatgcc	7260
ggcacgcgac	cgggcgcacc	gcagatggaa	acggccgacg	cgcagcttcg	cttcctctgc	7320
gaggcgggtt	tttcggccgg	ggacgccgtc	aatgcgctga	tgacaatcag	ctacttcact	7380

WO 2004/063359 PCT/EP2004/000099 11/357

gttggggccg	tgcttgagga	gcaggccggc	gacagcgatg	ccggcgagcg	cggcggcacc	7440
gttgaacagg	ctccgctctc	gccgctgttg	cgggccgcga	tagacgcctt	cgacgaagcc	7500
ggtccggacg	cagcgttcga	gcagggactc	gcggtgattg	tcgatggatt	ggcgaaaagg	7560
aggetegttg	tcaggaacgt	tgaaggaccg	agaaagggtg	acgattgatc	aggaccgctg	7620
ccggagegca	acccactcac	tacagcagag	ccatgtagac	aacatcccct	cecetttee	7680
accgcgtcag	acgcccgtag	cagcccgcta	cgggcttttt	catgccctgc	cctagcgtcc	7740
aagcctcacg	gccgcgctcg	gcctctctgg	cggccttctg	gegetettee	gcttcctcgc	7800
tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct	cactcaaagg	7860
cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	7920
gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggetee	7980
gcccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	8040
gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	cctgttccga	8100
ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgcttttcc	8160
gctgcataac	cctgcttcgg	ggtcattata	gcgattttt	cggtatatcc	atcctttttc	8220
gcacgatata	caggattttg	ccaaagggtt	cgtgtagact	ttccttggtg	tatccaacgg	8280
cgtcagccgg	gcaggatagg	tgaagtaggc	ccacccgcga	gcgggtgttc	cttcttcact	8340
gtcccttatt	cgcacctggc	ggtgctcaac	gggaatcctg	ctctgcgagg	ctggccggct	8400
accgccggcg	taacagatga	gggcaagcgg	atggctgatg	aaaccaagcc	aaccaggaag	8460
ggcagcccac	ctatcaaggt	gtactgcctt	ccagacgaac	gaagagcgat	tgaggaaaag	8520
geggeggegg	ccggcatgag	cctgtcggcc	tacctgctgg	ccgtcggcca	gggctacaaa	8580
atcacgggcg	tcgtggacta	tgagcacgtc	cgcgagctgg	cccgcatcaa	tggcgacctg	8640

WO 2004/063359 PCT/EP2004/000099 12/357

ggccgc	ctgg	gcggcctgct	gaaactctgg	ctcaccgacg	acccgcgcac	ggcgcggttc	8700
ggtgat	gcca	cgatcctcgc	cctgctggcg	aagatcgaag	agaagcagga	cgagcttggc	8760
aaggtc	atga	tgggcgtggt	ccgcccgagg	gcagagccat	gacttttta	gccgctaaaa	8820
cggccg	gggg	gtgcgcgtga	ttgccaagca	cgtccccatg	cgctccatca	agaagagcga	8880
cttcgc	ggag	ctggtgaagt	acatcaccga	cgagcaaggc	aagaccgagc	gcctttgcga	8940
cgctca	ccgg	gctggttgcc	ctcgccgctg	ggctggcggc	cgtctatggc	cctgcaaacg	9000
cgccag	aaac	gccgtcgaag	ccgtgtgcga	gacaccgcgg	ccgccggcgt	tgtggatacc	9060
tcgcgg	aaaa	cttggccctc	actgacagat	gaggggcgga	cgttgacact	tgaggggccg	9120
actcac	ccgg	cgcggcgttg	acagatgagg	ggcaggctcg	atttcggccg	gcgacgtgga	9180
gctggc	cagc	ctcgcaaatc	ggcgaaaacg	cctgatttta	cgcgagtttc	ccacagatga	9240
tgtgga	caag	cctggggata	agtgccctgc	ggtattgaca	cttgaggggc	gcgactactg	9300
acagat	gagg	ggcgcgatcc	ttgacacttg	aggggcagag	tgctgacaga	tgaggggcgc	9360
acctat	tgac	atttgagggg	ctgtccacag	gcagaaaatc	cagcatttgc	aagggtttcc	9420
gcccgt	tttt	cggccaccgc	taacctgtct	tttaacctgc.	ttttaaacca	atatttataa	9480
accttg	tttt	taaccagggc	tgcgccctgt	gcgcgtgacc	gcgcacgccg	aaggggggtg	9540
ccccc	cttc	tcgaaccctc	ccggcccgct	aacgcgggcc	teccatecee	ccaggggctg	9600
cgcccc	tcgg	ccgcgaacgg	cctcacccca	aaaatggcag	cgctggcagt	ccttgccatt	9660
gccggg	atcg	gggcagtaac	gggatgggcg	atcagcccga	gegegaegee	cggaagcatt	9720
gacgtg	ccgc	aggtgctggc	atcgacattc	agcgaccagg	tgccgggcag	tgagggcggc	9780
ggcctg	ggtg	geggeetgee	cttcacttcg	gccgtcgggg	cattcacgga	cttcatggcg	9840
gggccg	gcaa	tttttacctt	gggcattctt	ggcatagtgg	tcgcgggtgc	cgtgctcgtg	9900

WO 2004/063359 PCT/EP2004/000099 13/357

ttcgggggtg	cgataaaccc	agcgaaccat	ttgaggtgat	aggtaagatt	ataccgaggt	9960
atgaaaacga	gaattggacc	tttacagaat	tactctatga	agcgccatat	ttaaaaagct	10020
accaagacga	agaggatgaa	gaggatgagg	aggcagattg	ccttgaatat	attgacaata	10080
ctgataagat	aatatatctt	ttatatagaa	gatategeeg	tatgtaagga	tttcaggggg	10140
caaggcatag	gcagcgcgct	tatcaatata	tctatagaat	gggcaaagca	taaaaacttg	10200
catggactaa	tgcttgaaac	ccaggacaat	aaccttatag	cttgtaaatt	ctatcataat	10260
tgggtaatga	ctccaactta	ttgatagtgt	tttatgttca	gataatgccc	gatgactttg	10320
tcatgcagct	ccaccgattt	tgagaacgac	agcgacttcc	gtcccagccg	tgccaggtgc	10380
tgcctcagat	tcaggttatg	ccgctcaatt	cgctgcgtat	atcgcttgct	gattacgtgc	10440
agctttccct	tcaggcggga	ttcatacagc	ggccagccat	ccgtcatcca	tatcaccacg	10500
tcaaagggtg	acagcaggct	cataagacgc	cccagcgtcg	ccatagtgcg	ttcaccgaat	10560
acgtgcgcaa	caaccgtctt	ccggagactg	tcatacgcgt	aaaacagcca	gcgctggcgc	10620
gatttagccc	cgacatagcc	ccactgttcg	tccatttccg	cgcagacgat	gacgtcactg	10680
cccggctgta	tgcgcgaggt	taccgactgc	ggcctgagtt	ttttaagtga	cgtaaaatcg	10740
tgttgaggcc	aacgcccata	atgcgggctg	ttgcccggca	tccaacgcca	ttcatggcca	10800
tatcaatgat	tttctggtgc	gtaccgggtt	gagaagcggt	gtaagtgaac	tgcagttgcc	10860
atgttttacg	gcagtgagag	cagagatagc	gctgatgtcc	ggcggtgctt	ttgccgttac	10920
gcaccacccc	gtcagtagct	gaacaggagg	gacagctgat	agacacagaa	gccactggag	10980
cacctcaaaa	acaccatcat	acactaaatc	agtaagttgg	cagcatcacc	cataattgtg	11040
gtttcaaaat	cggctccgtc	gatactatgt	tatacgccaa	ctttgaaaac	aactttgaaa	11100
aagctgtttt	ctggtattta	aggttttaga	atgcaaggaa	cagtgaattg	gagttcgtct	11160

WO 2004/063359 PCT/EP2004/000099 14/357

tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 11220 aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11280 aaaagatacg gaaggaatgt ctcctgctaa ggtatataag ctggtgggag aaaatgaaaa 11340 cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11400 aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaaggtcc tgcactttga 11460 acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc tttgctcgga 11520 agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcatcag 11580 gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11640 · agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11700 agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11760 cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct ttgtgaaaga 11820 tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggtatga 11880 cattgccttc tgcgtccggt cgatcaggga ggatatcggg gaagaacagt atgtcgagct 11940 attttttgac ttactgggga tcaagcctga ttgggagaaa ataaaatatt atattttact 12000 ggatgaattg ttttagtacc tagatgtggc gcaacgatgc cggcgacaag caggagcgca 12060 ccgacttctt ccgcatcaag tgttttggct ctcaggccga ggcccacggc aagtatttgg 12120 gcaaggggtc gctggtattc gtgcagggca agattcggaa taccaagtac gagaaggacg 12180 gccagacggt ctacgggacc gacttcattg ccgataaggt ggattatctg gacaccaagg 12240 caccaggegg gtcaaatcag gaataagggc acattgeecc ggcgtgagtc ggggcaatcc 12300 cgcaaggagg gtgaatgaat cggacgtttg accggaaggc atacaggcaa gaactgatcg 12360 acgcggggtt ttccgccgag gatgccgaaa ccatcgcaag ccgcaccgtc atgcgtgcgc 12420

			15/35/			
cccgcgaaac	cttccagtcc	gtcggctcga	tggtccagca	agctacggcc	aagatcgagc	12480
gcgacagcgt	gcaactggct	caccatgada	tgcccgcgcc	atcggccgcc	gtggagcgtt	12540
cgcgtcgtct	cgaacaggag	gcggcaggtt	tggcgaagtc	gatgaccatc	gacacgcgag	12600
gaactatgac	gaccaagaag	cgaaaaaccg	ccggcgagga	cctggcaaaa	caggtcagcg	12660
aggccaagca	ggccgcgttg	ctgaaacaca	cgaagcagca	gatcaaggaa	atgcagcttt	12720
ccttgttcga	tattgcgccg	tggccggaca	cgatgcgagc	gatgccaaac	gacacggccc	12780
gctctgccct	gttcaccacg	cgcaacaaga	aaatcccgcg	cgaggcgctg	caaaacaagg	12840
tcattttcca	cgtcaacaag	gacgtgaaga	tcacctacac	cggcgtcgag	ctgcgggccg	12900
acgatgacga	actggtgtgg	cagcaggtgt	tggagtacgc	gaagcgcacc	cctatcggcg	12960
agccgatcac	cttcacgttc	tacgagcttt	gccaggacct	gggctggtcg	atcaatggcc	13020
ggtattacac	gaaggccgag	gaatgcctgt	cgcgcctaca	ggcgacggcg	atgggcttca	13080
cgtccgaccg	cgttgggcac	ctggaatcgg	tgtcgctgct	gcaccgcttc	cgcgtcctgg	13140
accgtggcaa	gaaaacgtcc	cgttgccagg	tcctgatcga	cgaggaaatc	gtcgtgctgt	13200
ttgctggcga	ccactacacg	aaattcatat	gggagaagta	ccgcaagctg	tegeegaegg	13260
cccgacggat	gttcgactat	ttcagctcgc	accgggagcc	gtacccgctc	aagctggaaa	13320
cetteegeet	catgtgcgga	tcggattcca	cccgcgtgaa	gaagtggcgc	gagcaggtcg	13380
gcgaagcctg	cgaagagttg	cgaggcagcg	gcctggtgga	acacgcctgg	gtcaatgatg	13440
acctggtgca	ttgcaaacgc	tagggccttg	tggggtcagt	teeggetggg	ggttcagcag	13500
ccagcgcttt	actggcattt	caggaacaag	cgggcactgc	tegaegeaet	tgcttcgctc	13560
agtatcgctc	gggacgcacg	gcgcgctcta	cgaactgccg	ataaacagag	gattaaaatt	13620
gacaattgtg	attaaggctc	agattcgacg	gcttggagcg	gccgacgtgc	aggatttccg	13680

WO 2004/063359 PCT/EP2004/000099 16/357

			10/35/			
cgagatccga	ttgtcggccc	tgaagaaagc	tccagagatg	ttcgggtccg	tttacgagca	13740
cgaggagaaa	aagcccatgg	aggegttege	tgaacggttg	cgagatgccg	tggcattcgg	13800
cgcctacatc	gacggcgaga	tcattgggct	gtcggtcttc	aaacaggagg	acggccccaa	13860
ggacgctcac	aaggcgcatc	tgtccggcgt	tttcgtggag	cccgaacagc	gaggccgagg	13920
ggţcgccggt	atgctgctgc	gggcgttgcc	ggcgggttta	ttgctcgtga	tgatcgtccg	13980
acagattcca	acgggaatct	ggtggatgcg	catcttcatc	ctcggcgcac	ttaatatttc	14040
gctattctgg	agcttgttgt	ttatttcggt	ctaccgcctg	ccgggcgggg	tcgcggcgac	14100
ggtaggcgct	gtgcagccgc	tgatggtcgt	gttcatctct	geegetetge	taggtagccc	14160
gatacgattg	atggcggtcc	tgggggctat	ttgcggaact	gcgggcgtgg	cgctgttggt	14220
gttgacacca	aacgcagcgc	tagatcctgt	cggcgtcgca	gcgggcctgg	cgggggcggt	14280
ttccatggcg	ttcggaaccg	tgctgacccg	caagtggcaa	cctcccgtgc	ctctgctcac	14340
ctttaccgcc	tggcaactgg	cggccggagg	acttctgctc	gttccagtag	ctttagtgtt	14400
tgatccgcca	atcccgatgc	ctacaggaac	caatgttctc	ggcctggcgt	ggctcggcct	14460
gatcggagcg	ggtttaacct	acttcctttg	gttccggggg	atctcgcgac	tcgaacctac	14520
agttgtttcc	ttactgggct	ttctcagccc	cagatctggg	gtcgatcagc	cggggatgca	14580
tcaggccgac	agtcggaact	tegggteece	gacctgtacc	attcggtgag	caatggatag	14640
gggagttgat	atcgtcaacg	ttcacttcta	aagaaatagc	gccactcagc	ttcctcagcg	14700
gctttatcca	gcgatttcct	attatgtcgg	catagttctc	aagatcgaca	gcctgtcacg	14760
gttaagcgag	aaatgaataa	gaaggctgat	aattcggatc	tctgcgaggg	agatgatatt	14820
tgatcacagg	cagcaacgct	ctgtcatcgt	tacaatcaac	atgctaccct	ccgcgagatc	14880
atccgtgttt	caaacccggc	agcttagttg	ccgttcttcc	gaatagcatc	ggtaacatga	14940

WO 2004/063359 PCT/EP2004/000099 17/357

gcaaagtctg ccgccttaca acggctctcc cgctgacgcc gtcccggact gatgggctgc 15000 ctgtatcgag tggtgatttt gtgccgagct gccggtcggg gagctgttgg ctggctggtg 15060 gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac acattgcgga 15120 cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa cagctgattg 15180 cccttcaccg cctggccctg agagagttgc agcaagcggt ccacgctggt ttgccccagc 15240 aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat aaatcaaaag 15300 aatageeega gatagggttg agtgttgtte eagtttggaa caagagteea etattaaaga 15360 acqtqqactc caacqtcaaa gggcgaaaaa ccgtctatca gggcgatggc ccactacgtg 15420 aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta aatcggaacc 15480 ctaaaqqqaq cccccqattt agaqcttgac ggggaaaqcc ggcgaacgtg gcgagaaagg 15540 aagggaagaa agcgaaagga gcgggcgcca ttcaggctgc gcaactgttg ggaagggcga 15600 teggtgeggg cetetteget attacgccag etggegaaag ggggatgtge tgeaaggega 15660 ttaagttggg taacgccagg gttttcccag tcacgacgtt gtaaaacgac ggccagtgaa 15720 15739 ttcgagctcg gtacccggg

<210> 4

<211> 11611

<212> DNA

<213> Artificial

<220>

<223> Vector

<220>

<221> misc_feature

<222> (227)..(227)

<223> n is a, c, g, or t

<220> <221> misc_feature <222> (318)..(318) <223> n is a, c, g, or t <220> <221> misc_feature <222>- (526)..(526) <223> n is a, c, g, or t <220> <221> misc_feature <222> (8946)..(8946) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10028)..(10028) <223> n is a, c, g, or t <400> 4 60 agettgeatg cetgeaggte gagtggagat gtggagtggg cgettacaca gtacacgagg acttctagct agaaagaagg attacctcta aacaagtgta cctgtgcatt ctgggtaaac 120 gactcatagg agagttgtaa aaaagtttcg gccggcgtat tgggtgttac ggagcattca 180 ctaggcaacc atgcatggtt actattgtat accatcttag taggaantga tttcgaggtt 240 300 tatacctacg atgaatgtgt gtcctgtagg cttgagagtt caaggaagaa acatgcaatt 360 atctttgcga acccaggngc tggtgacgga attttcatag tcaagctatc agagtaaaga agaggagcat gtcaaagtac aattagagac aaatatatag tcgcgtggag ccaagagcgg 420 attectcagt ctcgtaggtc tcttgacgac cgttgatctg cttgatctcg tctcccgaaa 480 540 atgaaaatag ctctgctaag ctattcttct cttcgccgga gcctgnaagg cgttactagg

ttgcagtcaa tgcattaatg cattgcagat gagctgtatc tggaagaggt aaacccgaaa

600

WO 2004/063359 PCT/EP2004/000099 19/357

acgcgtttta	ttcttgttga	catggagcta	ttaaatcact	agaaggcact	ctttgctgct	660
tggacaaatg	aacgtatctt	atcgagatcc	tgaacaccat	ttgtctcaac	tccggagctg	720
acatcgacac	caacgatctt	atatccagat	tegteaaget	gtttgatgat	ttcagtaacg	780
ttaagtggat	cgatcccgcg	gtcggcatct	actctattcc	tttgccctcg	gacgagtgct	840
ggggcgtcgg	tttccactat	cggcgagtac	ttctacacag	ccatcggtcc	agacggccgc	900
gcttctgcgg	gcgatttgtg	tacgcccgac	agtcccggct	ccggatcgga	cgattgcgtc	960
gcatcgaccc	tgcgcccaag	ctgcatcatc	gaaattgccg	tcaaccaagc	tctgatagag	1020
ttggtcaaga	ccaatgcgga	gcatatacgc	ccggagccgc	ggcgatcctg	caagctccgģ	1080
atgecteege	tcgaagtagc	gcgtctgctg	ctccatacaa	gccaaccacg	gcctccagaa	1140
gaagatgttg	gegaeetegt	attgggaatc	cccgaacatc	gcctcgctcc	agtcaatgac	1200
cgctgttatg	cggccattgt	ccgtcaggac	attgttggag	ccgaaatccg	cgtgcacgag	1260
gtgccggact	teggggeagt	cctcggccca	aagcatcagc	tcatcgagag	cctgcgcgac	1320
ggacgcactg	acggtgtcgt	ccatcacagt	ttgccagtga	tacacatggg	gatcagcaat	1380
cgcgcatatg	aaatcacgcc	atgtagtgta	ttgaccgatt	ccttgcggtc	cgaatgggcc	1440
gaacccgctc	gtctggctaa	gatcggccgc	agcgatcgca	tccatggcct	ccgcgaccgg	1500
ctgcagaaca	gcgggcagtt	cggtttcagg	caggtcttgc	aacgtgacac	cctgtgcacg	1560
gcgggagatg	caataggtca	ggeteteget	gaattcccca	atgtcaagca	cttccggaat	1620
cgggagcgcg	gccgatgcaa	agtgccgata	aacataacga	tctttgtaga	aaccatcggc	1680
gcagctattt	acccgcagga	catatccacg	ccctcctaca	tcgaagctga	aagcacgaga	1740
ttattagaca	tccgagagct	gcatcaggtc	ggagacgctg	tcgaactttt	cgatcagaaa	1800
cttctcgaca	gacgtcgcgg	tgagttcagg	catggtgatg	tctgctcaag	cggggtagct	1860

WO 2004/063359 PCT/EP2004/000099 20/357

			20,20,			
gttagtcaag	ctgcgatgaa	gtgggaaagc	tcgaactgaa	aggttcaaag	gaataaggga	1920
tgggaaggat	ggagtatgga	tgtagcaaag	tacttactta	ggggaaataa	aggttcttgg	1980
atgggaagat	gaatatactg	aagatgggaa	aagaaagaga	aaagaaaaga	gcagctggtg	2040
gggagagcag	gaaaatatgg	caacaaatgt	tggactgacg	caacgacctt	gtcaaccccg	2100
ccgacacacc	gggcggacag	acggggcaaa	gctgcctacc	agggactgag	ggacctcagc	2160
aggtcgagtg	cagagcaccg	gatgggtcga	ctgccagctt	gtgttcccgg	tetgegeege	2220
tggccagctc	ctgagcggcc	tttccggttt	catacaccgg	gcaaagcagg	agaggcacga	2280
tatttggacg	ccctacagat	gccggatggg	ccaattaggg	agcttacgcg	ccgggtactc	2340
gctctaccta	cttcggagaa	ggtactatct	cgtgaatctt	ttaccagatc	ggaagcaatt	2400
ggacttctgt	acctaggtta	atggcatgct	atttcgccga	cggctataca	cccctggctt	2460
cacattctcc	ttcgcttact	gccggtgatt	cgatgaagct	ccatattctc	cgatgatgca	2520
atagattctt	ggtcaacgag	gggcacacca	gcctttccac	ttcggggcgg	aggggcggcc	2580
ggtcccggat	taataatcat	ccactgcacc	tcagageege	cagagetgte	tggccagtgg	2640
cttattactc	agcccttctc	tctgcgtccg	teegtetete	cgcatgccag	aaagagtcac	2700
cggtcactgt	acagagetea	cgagttcgtc	acatttttct	acaaatggtg	gaggcggcgg	2760
attttaggct	caagtcatga	ccctctgggt	cactccagaa	tcagctaggt	caacgaataa	2820
ggatgattct	ataggaagat	ccaggcaccg	gtcaaccatg	atctggacag	atttgggagc	2880
tcggtataag	ctctccacct	atcttattct	gtatagttta	ggcttaaagt	ttatccagga	2940
gatgttgctg	aagtcgattt	gagtccactt	cctcactggt	agctatacga	ctttgatggt	3000
cgttgtaggg	gctgtattag	gtctcgatca	aacacaaata	gaattaaatg	gtactcgagt	3060
ccactgaagg	tggcttctcc	gtcttccgta	gccgtgccga	aatccttaca	gcttgtgttg	3120

WO 2004/063359 PCT/EP2004/000099 21/357

tgtgactttt	ggttacgccg	tctgactttt	gtggtgagct	aactagagat	catgctatat	3180
ctcctgattt	aatacaatgc	tcatcataac	attccacctg	gaactgctag	caacgtttga	3240
cttgcattgt	gcaacgccct	ttgcagagct	atcggatgat	caatagtgcc	acgttctaaa	3300
ttcaaccaac	gcaggtgccc	caagcetteg	acatccggat	gtatttcgaa	aacctcatgg	3360
cgattgcagt	cctcagattc	atgttcattc	caatgctcat	tggtgaataa	aaggttcaca	3420
gggaataagt	tcaaactcga	gatacttgag	aatattgaaa	gccaaaggac	cctctatgct	3480
ccaagctaga	gtctcagcct	ggaaagcaaa	tccaaatgaa	gctatgctac	ctccaattcc	3540
tcatcatctt	atctataata	cagagtcgaa	gaatatcctc	ttgacaccgc	teegteetee	3600
gacttcaata	aggagcttac	tcctccttga	caccacccct	ccagttcttc	teggegttet	3660
ggagggaggc	cttgtcggtc	ttgggctggc	cctggctgag	aaagctgttg	gcagccttaa	3720
agggacgctg	gaggtcacca	gtcgctggct	tcccgaagac	gtggatctta	accagattcg	3780
aaagcgcctt	cagcggatga	tcgactggat	cagaagagcg	ttggtgtact	tgaagtacag	3840
atgcatgacg	gccatcatgc	caacgcccat	gaactggctc	ttaatgagct	ggcggaactg	3900
gcccttatcg	tactccatgt	tggtagttgt	gacaggacga	ggctcctcgc	cgcttccaag	3960
cggagcaggc	tcgacgtatt	tcagtgtcga	aagatctgat	caagagacag	gatgaggatc	4020
gtttcgcatg	attgaacaag	atggattgca	cgcaggttct	ccggccgctt	gggtggagag	4080
gctattcggc	tatgactggg	cacaacagac	aatcggctgc	tctgatgccg	ccgtgttccg	4140
gctgtcagcg	caggggcgcc	cggttctttt	tgtcaagacc	gacctgtccg	gtgccctgaa	4200
tgaactgcag	gacgaggcag	cgcggctatc	gtggctggcc	acgacgggcg	ttccttgcgc	4260
agctgtgctc	gacgttgtca	ctgaagcggg	aagggactgg	ctgctattgg	gcgaagtgcc	4320
ggggcaggat	ctcctgtcat	ctcaccttgc	tcctgccgag	aaagtatcca	tcatggctga	4380

WO 2004/063359 PCT/EP2004/000099 22/357

tgcaatgcgg	cggctgcata	cgcttgatcc	ggctacctgc	ccattcgacc	accaagcgaa	4440
acategeate	gagcgagcac	gtactcggat	ggaagccggt	cttgtcgatc	aggatgatct	4 500
ggacgaagag	catcaggggc	tegegeeage	cgaactgttc	gccaggctca	aggcgcgcat	4560
gcccgacggc	gaggateteg	tcgtgaccca	tggcgatgcc	tgcttgccga	atatcatggt	4620
ggaaaatggc	cgcttttctg	gattcatcga	ctgtggccgg	ctgggtgtgg	cggaccgcta	4680
tcaggacata	gcgttggcta	cccgtgatat	tgctgaagag	cttggcggcg	aatgggctga	4740
cegetteete	gtgctttacg	gtatcgccgc	tcccgattcg	cagcgcatcg	ccttctatcg	4800
ccttcttgac	gagttcttct	gagcgggact	ctggggttcg	aaatgaccga	ccaagcgacg	4860
cccaacctgc	catcacgaga	tttcgattcc	accgccgcct	tctatgaaag	gttgggcttc	4920
ggaatcgttt	tccgggacgc	cggctggatg	atcctccagc	gcggggatct	catgctggag	498 <u>0</u>
ttcttcgccc	accccgggct	cgatcccctc	gcgagttggt	tcagctgctg	cctgaggctg	5040
gacgacctcg	cggagttcta	ccggcagtgc	aaatccgtcg	gcatccagga	aaccagcagc	5100
ggctatccgc	gcatccatgc	ccccgaactg	caggagtggg	gaggcacgat	ggccgctttg	5160
gtccggatct	ttgtgaagga	accttacttc	tgtggtgtga	cataattgga	caaactacct	5220
acagagattt	aaagctctaa	ggtaaatata	aaatttttaa	gtgtataatg	tgttaaacta	5280
ctgattctaa	ttgtttgtgt	attttagatt	ccaacctatg	gaactgatga	atgggagcag	5340
tggtggaatg	cctttaatga	ggaaaacctg	ttttgctcag	aagaaatgcc	atctagtgat	5400
gatgaggcta	ctgctgactc	tcaacattct	actcctccaa	aaaagaagag	aaaggtagaa	5460
gaccccaagg	actttccttc	agaattgcta	agttttttga	gtcatgctgt	gtttagtaat	5520
agaactcttg	cttgctttgc	tatttacacc	acaaaggaaa	aagctgcact	gctatacaag	5580
aaaattatgg	aaaaatattc	tgtaaccttt	ataagtaggc	ataacagtta	taatcataac	5640

atactgtttt	ttcttactcc	acacaggcat	agagtgtctg	ctattaataa	ctatgctcaa	5700
aaattgtgta	cctttagctt	tttaatttgt	aaaggggtta	ataaggaata	tttgatgtat	5760
agtgccttga	ctagagatca	taatcagcca	taccacattt	gtagaggttt	tacttgcttt	5820
aaaaaacctc	ccacacctcc	ccctgaacct	gaaacataaa	atgaatgcaa	ttgttgttgt	5880
taacttgttt	attgcagctt	ataatggtta	caaataaagc	aatagcatca	caaatttcac	5940
aaataaagca	ttttttcac	tgcattctag	ttgtggtttg	tccaaactca	tcaatgtatc	6000
ttatcatgtc	tggatctgac	gggtgcgcat	gatcgtgctc	ctgtcgttga	ggacccggct	6060
aggctggcgg	ggttgcctta	ctggttagca	gaatgaatca	ccgatacgcg	agcgaacgtg	6120
aagcgactgc	tgctgcaaaa	cgtctgcgac	ctgagcaaca	acatgaatgg	tcttcggttt	6180
ccgtgtttcg	taaagtctgg	aaacgcggaa	gtcagcgctc	ttccgcttcc	tcgctcactg	6240
actcgctgcg	ctcggtcgtt	cggctgcggc	gagcggtatc	agctcactca		6300
tacggttatc	cacagaatca	ggggataacg	caggaaagaa	catgtgagca	aaaggccagc	6360
aaaaggccag	caaaaggcca	ggaaccgtaa	aaaggccgcg	ttgctggcgt	ttttccatag	6420
gctccgcccc	cctgacgagc	atcacaaaaa	tcgacgctca	agtcagaggt	ggcgaaaccc	6480
gacaggacta	taaagatacc	aggcgtttcc	ccctggaagc	tccctcgtgc	gctctcctgt	6540
tccgaccctg	ccgcttaccg	gatacctgtc	cgcctttctc	ccttcgggaa	gcgtggcgct	6600
ttctcatagc	tcacgctgta	ggtatctcag	ttcggtgtag	gtcgttcgct	ccaagctggg	6660
ctgtgtgcac	gaaccccccg	ttcagcccga	ccgctgcgcc	ttatccggta	actatcgtct	6720
tgagtccaac	ccggtaagac	acgacttatc	gccactggca	gcagccactg	gtaacaggat	6780
tagcagagcg	aggtatgtag	gcggtgctac	agagttcttg	aagtggtggc	ctaactacgg	6840

WO 2004/063359 PCT/EP2004/000099 24/357

	ctacactaga	aggacagtat	ttggtatctg	cgctctgctg	aagccagtta	ccttcggaaa	6900 .
•	aagagttggt	agctcttgat	ccggcaaaca	aaccaccgct	ggtagcggtg	gtttttttgt	6960
	ttgcaagcag	cagattacgc	gcagaaaaaa	aggatctcaa	gaagatcctt	tgatcttttc	7020
	tacggggtct	gacgctcagt	ggaacgaaaa	ctcacgttaa	gggattttgg	tcatgagatt	7080
	atcaaaaagg	atcttcacct	agatcctttt	aaattaaaaa	tgaagtttta	aatcaatcta	7140
	aagtatatat	gagtaaactt	ggtctgacag	ttaccaatgc	ttaatcagtg	aggcacctat	7200
	ctcagcgatc	tgtctatttc	gttcatccat	agttgcctga	ctccccgtcg	tgtagataac	7260
	tacgatacgg	gagggcttac	catctggccc	cagtgctgca	atgataccgc	gagacccacg	7320
	ctcaccggct	ccagatttat	cagcaataaa	ccagccagcc	ggaagggccg	agcgcagaag	7380
	tggtcctgca	actttatccg	cctccatcca	gtctattaat	tgttgccggg	aagctagagt	7440
	aagtagttcg	ccagttaata	gtttgcgcaa	cgttgttgcc	attgctgcag	gcatcgtggt	7500
	gtcacgctcg	tcgtttggta	tggcttcatt	cageteeggt	tcccaacgat	caaggcgagt	7560
	tacatgatcc	cccatgttgt	gcaaaaaagc	ggttagctcc	tteggteete	cgatcgttgt	7620
	cagaagtaag	ttggccgcag	tgttatcact	catggttatg	gcagcactgc	ataattctct	7680
	tactgtcatg	ccatccgtaa	gatgcttttc	tgtgactggt	gagtactcaa	ccaagtcatt	7740
	ctgagaatag	tgtatgcggc	gaccgagttg	ctcttgcccg	gcgtcaacac	gggataatac	7800
	cgcgccacat	agcagaactt	taaaagtgct	catcattgga	aaacgttctt	cggggcgaaa	7860
	acteteaagg	atcttaccgc	tgttgagatc	cagttcgatg	taacccactc	gtgcacccaa	7920
	ctgatcttca	gcatctttta	ctttcaccag	cgtttctggg	tgagcaaaaa	caggaaggca	7980
	aaatgccgca	aaaaagggaa	taagggcgac	acggaaatgt	tgaatactca	tactcttcct	8040
	ttttcaatat	tattgaagca	tttatcaggg	ttattgtctc	atgagcggat	acatatttga	8100

WO 2004/063359 PCT/EP2004/000099 25/357

atgtatttag	aaaaataaac	aaataggggt	tccgcgcaca	tttccccgaa	aagtgccacc	8160
tgacgtctaa	gaaaccatta	ttatcatgac	attaacctat	aaaaataggc	gtatcacgag	8220
gccctttcgt	cttcaagaat	tegeggeege	aattaaccct	cactaaagga	tccctatagt	8280
gagtcgtatt	atgeggeege	gaattctcat	gtttgaccgc	ttatcatcga	taagctctgc	8340
tttttgttga	cttccattgt	tcattccacg	gacaaaaaca	gagaaaggaa	acgacagagg	8400
ccaaaaagct	cgctttcagc	acctgtcgtt	tectttettt	tcagagggta	ttttaaataa	8460
aaacattaag	ttatgacgaa	gaagaacgga	aacgccttaa	accggaaaat	tttcataaat	8520
agcgaaaacc	cgcgaggtcg	ccgccccgta	acaaggcgga	tcgccggaaa	ggacccgcaa	8580
atgataataa	ttatcaattg	catactatcg	acggcactgc	tgccagataa	caccaccggg	8640
gaaacattcc	atcatgatgg	ccgtgcggac	ataggaagcc	agttcatcca	tcgctttctt	8700
gtctgctgcc	atttgctttg	tgacatccag	cgccgcacat	tcagcagcgt	ttttcagcgc	8760
gttttcgatc	aacgtttcaa	tgttggtatc	aacaccaggt	ttaactttga	acttatcggc	8820
actgacggtt	accttgttct	gcgctggctc	atcacgcagg	ataccaaggc	tgatgttgta	8880
gatattggtc	accggctgag	ggttttcgat	tgccgctgcg	tggatagcac	catttgcgat	8940
caggcngtcc	ttgatgaatg	acactccatt	gcgaataagt	tcgaaggaga	cggtgtcacg	9000
aatgcgctgg	tccagctcgg	tcgattgcct	tttgtgcagc	agaggtatca	atctcaacgc	9060
caaggctcat	cgaagcgcaa	tattgctgct	caccaaaacg	cgtattgacc	aggtgttcaa	9120
cggcaaattt	ctgcccttct	gatgtcagaa	aggcaaagtg	attttctttc	tggtattcag	9180
ttgctgtgtg	tcggtttcag	caaaaccaag	ctcgcgcaat	teggetgtge	agatttagaa	9240
ggcagatcac	cagacagcaa	cggccaacgg	aaaacagcgc	atacagaaca	teegtegeeg	9300
cgccgacaac	gtgataattt	ttatgaccca	tgatttattt	ccttttagac	gtgagcctgt	9360

WO 2004/063359 PCT/EP2004/000099 26/357

				20/35/			
cgc	acagcaa	agccgccgaa	agttcctcga	agctagcttc	agacgtgtct	agatacgtct	9420
gct	ttttgtt	gacttccatt	gttcattcca	cggacaaaaa	cagagaaagg	aaacgacaga	9480
ggc	caaaaag	ctcgctttca	gcacctgtcg	tttcctttct	tttcagaggg	tattttaaat	9540
aaa	aacatta	agttatgacg	aagaagaacg	gaaacgcctt	aaaccggaaa	attttcataa	9600
ata	gcgaaaa	cccgcgaggt	cgccgccccg	taacaaggcg	gategeegga	aaggaccege	9660
aaa	tgataat	aattatcaat	tgcatactat	cgacggcact	gctgccagat	aacaccaccg	9720
ggg	aaacatt	ccatcatgat	ggccgtgcgg	acataggaag	ccagttcatc	catcgctttc	9780
ttg	tctgctg	ccatttgctt	tgtgacatcc	agcgccgcac	attcagcagc	gtttttcagc	9840
gcg	ttttcga	tcaacgtttc	aatgttggta	tcaacaccag	gtttaacttt	gaacttatcg	9900
gca	ctgacgg	ttaccttgtt	ctgcgctggc	tcatcacgca	ggataccaag	gctgatgttg	9960
tag	atattgg	tcaccggctg	agggttttcg	attgccgctg	cgtggatagc	accatttgcg	10020
atc	aggcngt	ccttgatgaa	tgacactcca	ttgcgaataa	gttcgaagga	gacggtgtca	10080
cga	atgcgct	ggtccagctc	ggtcgattgc	cttttgtgca	gcagaggtat	caatctcaac	10140
gcc	aaggctc	atcgaagcgc	aatattgctg	ctcaccaaaa	cgcgtattga ·	ccaggtgttc	10200
aac	ggcaaat	ttctgccctt	ctgatgtcag	aaaggcaaag	tgattttctt	tctggtattc	10260
agt	tgctgtg	tgtcggtttc	agcaaaacca	agctcgcgca	attcggctgt	gcagatttag	10320
aag	gcagatc	accagacagc	aacggccaac	ggaaaacagc	gcatacagaa	cateegtege	10380
cgc	gccgaca	acgtgataat	ttttatgacc	catgatttat	ttccttttag	acgtgagcct	10440
gtc	gcacagc	aaagccgccg	aaagttcctc	gaccgatgcc	cttgagagcc	ttcaacccag	10500
tca	gctcctt	ccggtgggcg	cggggcatga	ctatcgtcgc	cgcacttatg	actgtcttct	10560
tta	tcatgca	actcgtagga	caggtgccgg	cagcgctctg	ggtcattttc	ggcgaggacc	10620

WO 2004/063359 PCT/EP2004/000099 27/357

gctttcgctg gagcgcgacg atgatcggcc tgtcgcttgc ggtattcgga atcttgcacg 10680 ccctcgctca agccttcgtc actggtcccg ccaccaaacg tttcggcgag aagcaggcca 10740 ttatcgccgg catggcggcc gacgcgttgg gctacgtctt gctggcgttc gcgacgcgag 10800 gctggatggc cttccccatt atgattcttc tcgcttccgg cggcatcggg atgcccgcgt 10860 tgcaggccat gctgtccagg caggtagatg acgaccatca gggacagctt caaggatcgc 10920 tegeggetet taccageeta acttegatea ttggaceget gategteaeg gegatttatg 10980 ccgcctcggc gagcacatgg aacgggttgg catggattgt aggcgccgcc ctataccttg 11040 tetgeeteee egegttgegt egeggtgeat ggageeggge eacetegace tgaatggaag 11100 ccggcggcac ctcgctaacg gattcaccac tccaagaatt ggagccaatc aattcttgcg 11160 gagaactgtg aatgcgcaaa ccaaccttg gcagaacata tccatcgcgt ccgccatctc 11220 cagcagccgc acgcggcgca tctcgggcag cgttgggtcc tgcagatccg gctgtggaat 11280 gtgtgtcagt tagggtgtgg aaagtcccca ggctccccag caggcagaag tatgcaaagc 11340 atgcatctca attagtcagc aaccaggtgt ggaaagtccc caggctcccc agcaggcaga 11400 agtatgcaaa gcatgcatct caattagtca gcaaccatag tcccgcccct aactccgccc 11460 atcocgccc taactccgcc cagttccgcc cattctccgc cccatggctg actaattttt 11520 tttatttatg cagaggccga ggccgcctcg gcctctgagc tattccagaa gtagtgagga 11580 11611 ggcttttttg gaggcctagg cttttgcaaa a

<210> 5

<211> 21

<212> DNA

<213> Artificial

•	WO 2004/063359	PCT/EP2004/ 28/357	000099
<223>	Primer		•
<400>	5		
cgatg	tagga gggcgtggat	a .	21
<210>	6		
<211>	21		
<212>			
<213>	Artificial		
		·	
<220>			
<223>	Primer		
.400	_		
<400>			0.1
gette	tgcgg gcgatttgtg	t	21
<210>	7		
<211>			
<212>			
	Artificial		
<220>			
<223>	Primer		
<400>	7		
tgagaa	atatc accggaattg	:	20
<210>	8	·	
<211>	21		
<212>	DNA		
<213>	Artificial	•	
<220>			
<223>	Primer		
	- 		
<400>	8		

21

agctcgacat actgttcttc c

v	/O 2004/063359	29/357		PCT/EP200	4/000099
	9				
	24				
.>	DNA				
	Artificial				
20,		•			
· <220>					
<223>	Primer				
<400>	9			•	
gtgaat	ggaa atcccatcgc tgtc				24
<210>	10				
<211>	24 .				
<212>	DNA		•		
<213>	Artificial		•		
<220>					
<223>	Primer				
<400>					
agtggg	tact ctaaaggcca tacc				24
<210>	11				
<211>	1771		•		
<212>	DNA				
<213>					
1220	nacination praviation				
<220>			•		
<221>	CDS				
<222>	(166)(1155)				
<400>	11				
ggcacg	aget tgeacgeaag teagegegeg e	aagtcaaca	cctgccggtc	cacagcctca .	60

aataataaag ageteaageg tttgtgegee tegaegtgge eagtetgeae tgeettgaae

ccgcgagtct cccgccgcac tgactgccat agcacagcta gacga atg cag cta gca

120

177

Met Gln Leu Ala

1

gcg aca gta atg ttg gag cag ctt acc gga agc gct gag gca ctc aag 225 Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala Glu Ala Leu Lys .10 15 gag aag gag gag gtt gca ggc agc tct gac gtg ttg cgt aca tgg 273 Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val Leu Arg Thr Trp 25 30 geg acc cag tac teg ett eeg tea gaa gag tea gae geg gee ege eeg 321. Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp Ala Ala Arg Pro 40 45 50 gga ctg aag aat gcc tac aag cca cct tcc gac aca aag ggc atc 369 . Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp Thr Lys Gly Ile 55 60 65 aca atg gcg cta cgt gtc atc ggc tcc tgg gcc gca gtg ttc ctc cac . 417 Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala Val Phe Leu His 70 75 80 gcc att ttt caa atc aag ctt ccg acc tcc ttg gac cag ctg cac tgg 465 Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp Gln Leu His Trp 85 90 95 100 ctg ccc gtg tca gat gcc aca gct cag ctg gtt agc ggc acg agc agc 513 Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser Gly Thr Ser Ser 105 110 115 ctg ctc gac atc gtc gta gta ttc ttt gtc ctg gag ttc ctg tac aca 561 Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu Phe Leu Tyr Thr 120 125 130 ggc ctt ttt atc acc acg cat gat gct atg cat ggc acc atc gcc atg 609 Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly Thr Ile Ala Met 135 · 140 145 aga aac agg cag ctt aat gac ttc ttg ggc aga gta tgc atc tcc ttg 657 Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val Cys Ile Ser Leu 150 155 160

									31	1331						
tac	gcc	tgg	ttt	gat	tac	aac	atg	ctg	cac	cgc	aag	cat	tgg	gag	cac	705
Tyr	Ala	Trp	Phe	Asp	Tyr	Asn	Met	Leu	His	Arg	Lys	His	Trp	Glu	His	
165					170					175					180	
					٠											
cac	aac	cac	act	ggc	gag	gtg	ggc	aag	gac	cct	gac	ttc	cac	agg	gga	753
												Phe				
-				185			,		190					195		
220	cct	aac	att	ata	ccc	taa	ttt	acc	agc	ttc	atg	tcc	agc	tac	atg	801
												Ser				
ASII	FIO	GLY	200	vul	110		-110	205					210			
			200					205								
		.					ata	~~~	taa	taa	200	gtg	ata	ato	car	849
_																0.25
Ser	Met		GIN	Pne	Ala	Arg		Ala	тър	ırb	TIIT	Val	vaı	Mec	GIII	
		215					220					225			•	
																007
												atg				897
Leu	Leu	Gly	Ala	Pro	Met	Ala	Asn	Leu	Leu	Val		Met	Ala	Ala	Ala	
	230					235					240					
												acg				945
Pro	Ile	Leu	Ser	Ala	Phe	Arg	Leu	Phe	Tyr	Phe	Gly	Thr	Tyr	Met	Pro	
245					250					255					260	
cac	aag	cct	gag	cct	ggc	gcc	gcg	tca	ggc	tct	tca	cca	gcc	gtc	atg	993
His	Lys	Pro	Glu	Pro	Gly	Ala	Ala	Ser	Gly	Ser	Ser	Pro	Ala	Val	Met	
				265					270					275		
										•			•			
aac	tgg	tgg	aag	tcg	cgc	act	agc	cag	gcg	tcc	gac	ctg	gtc	agc	ttt	1041
												Leu				
	_		280					285					290			
cta	acc	tac	tac	cac	ttc	gac	ctg	cac	tgg	gag	cac	cac	cgc	tgg	ccc	1089
															Pro	
Dou	****	295					300		-			305				
		2,7,5					500									
++~	acc	ccc	taa	taa	gad	ata	dac	aac	tac	cac	cac	cta	tct	ggc	cga	1137
															Arg	
rne			ττb	rrp	GIU			.1011	-13	9	320			1	3	
	310					315					220					•
						_1		a na	+~~-	at ~~	· « «	ataa	taca	a		1185
						ctg	gaca	cac	Lyca	gugg	ge e	ctgc	Lycc	·u		1100
Gly	Leu	. Val	Pro	Ala	L											

WO 2004/063359 PCT/EP2004/000099 32/357

325

gctgggcatg caggttgtgg caggactggg tgaggtgaaa agctgcaggc gctgctgccg gacacgctgc atgggctacc ctgtgtagct gccgccacta ggggaggggg tttgtagctg 1305 togagottgo cocatggatg aagotigtgta gtggtgcagg gagtacacco acaggccaac 1425 accettgeag gagatgtett gegtegggag gagtgttggg cagtgtagat getatgattg tatettaatg ctgaageett taggggageg acaettagtg ctgggcagge aacgeeetge 1485 aaggtgcagg cacaagctag gctggacgag gactcggtgg caggcaggtg aagaggtgcg 1545 ggagggtggt gccacaccca ctgggcaaga ccatgctgca atgctggcgg tgtggcagtg 1605 agagctgcgt gattaactgg gctatggatt gtttgagcag tctcacttat tctttgatat 1665 agatactggt caggcaggtc aggagagtga gtatgaacaa gttgagaggt ggtgcgctgc 1725 1771 ccctgcgctt atgaagctgt aacaataaag tggttcaaaa aaaaaa

<210> 12

<211> 329

<212> PRT

<213> Haematococcus pluvialis

<400> 12

Met Gln Leu Ala Ala Thr Val Met Leu Glu Gln Leu Thr Gly Ser Ala 1 5 10 15

Glu Ala Leu Lys Glu Lys Glu Lys Glu Val Ala Gly Ser Ser Asp Val 20 25 30

Leu Arg Thr Trp Ala Thr Gln Tyr Ser Leu Pro Ser Glu Glu Ser Asp
35 40 45

Ala Ala Arg Pro Gly Leu Lys Asn Ala Tyr Lys Pro Pro Pro Ser Asp 50 . 55 60

Thr Lys Gly Ile Thr Met Ala Leu Arg Val Ile Gly Ser Trp Ala Ala 65 70 75 80

Val Phe Leu His Ala Ile Phe Gln Ile Lys Leu Pro Thr Ser Leu Asp 85 90 95

Gln Leu His Trp Leu Pro Val Ser Asp Ala Thr Ala Gln Leu Val Ser 100 105 110 '

Gly Thr Ser Ser Leu Leu Asp Ile Val Val Val Phe Phe Val Leu Glu
115 120 125

Phe Leu Tyr Thr Gly Leu Phe Ile Thr Thr His Asp Ala Met His Gly .
130 135 140

Thr Ile Ala Met Arg Asn Arg Gln Leu Asn Asp Phe Leu Gly Arg Val

Cys Ile Ser Leu Tyr Ala Trp Phe Asp Tyr Asn Met Leu His Arg Lys
165 170 170 175

His Trp Glu His His Asn His Thr Gly Glu Val Gly Lys Asp Pro Asp 180 . 185 . 190

Phe His Arg Gly Asn Pro Gly Ile Val Pro Trp Phe Ala Ser Phe Met
195 200 205

Ser Ser Tyr Met Ser Met Trp Gln Phe Ala Arg Leu Ala Trp Trp Thr

210 215 220

Val Val Met Gln Leu Leu Gly Ala Pro Met Ala Asn Leu Leu Val Phe
225 230 235 240

Met Ala Ala Ala Pro Ile Leu Ser Ala Phe Arg Leu Phe Tyr Phe Gly
245 250 255

Thr Tyr Met Pro His Lys Pro Glu Pro Gly Ala Ala Ser Gly Ser Ser 260 265 270

Pro Ala Val Met Asn Trp Trp Lys Ser Arg Thr Ser Gln Ala Ser Asp 275 280 285

Leu Val Ser Phe Leu Thr Cys Tyr His Phe Asp Leu His Trp Glu His 290 295 300

His Arg Trp Pro Phe Ala Pro Trp Trp Glu Leu Pro Asn Cys Arg Arg 305 310 315 320

· Leu Ser Gly Arg Gly Leu Val Pro Ala 325

<210> 13

فتخر.

<211> 1662

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (168)..(1130)

<400> 13

									35/	357						
cgg:	ggca	act	caag	aaat	tc a	acago	ctgc	a ag	cgcg	cccc	agc	ctca	cag	cgcc	aagtga	60
gcta	atcg	acg	tggt	tgtg	ag c	gctc	gacg	t gg	tcca	ctga	cgg	gaat	gtg	agcc	tctgcg	120
ctc	egte	ctc ·	tgcca	aaat	ct c	gegto	eggg	g cc	tgcc	taag	tcg	aaga		cac		176
_												_		gct Ala		224
														cca Pro		272
	_		_	_	_	_				_		_		aaa Lys 50		320
	_		_	_	_			_						att Ile		368
														cta Leu	_	416
			-	_										aca Thr	_	464
		_			_	_	_		_			_	-	gtc Val		512
	_				_									cat His 130	_	560
gca	atg	cat	ggc	acc	ata	gct	ttg -	agg	cac	agg	cag	ctc	aat -	gat -	ctc	608

Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu Asn Asp Leu

135 140 145

att	ממכ	220	ato	tac	a+ a	tan	a+~			4					_4	CF.C
															atg	656
пеа	СТУ		TTE	Cys	тте	ser		туг	Ala	urp	Pne	Asp	луг	ser	Met	
		150					155					160				
+					.											
												ggc				704
Гел		Arg	Lys	His	Trp		His	His	Asn	His	Thr	Gly	Glu	Val	Gly	
	165					170					175					
												gtc				752
Lys	Asp	Pro	Asp	Phe	His	Lys	Gly	Asn	Pro	Gly	Leu	Val	Pro	Trp	Phe	
180					185					190					195	
												ttt				800
Ala	Ser	Phe	Met	Ser	Ser	Tyr	Met	Ser	Leu	Trp	Gln	Phe	Ala	Arg	Leu	
				200					205					210		
gca	tgg	tgg	gca	gtg	gtg	atg	caa	atg	ctg	ggg	gcg	ccc	atg	gca	aat	848
Ala	Trp	Trp	Ala	Val	Val	Met	Gln	Met	Leu	Gly	Ala	Pro	Met	Ala	Asn	
			215					220					225			
ctc	cta	gtc	ttc	atg	gct	gca	gcc	cca	atc	ttg	tca	gca	ttc	cgc	ctc	896
Leu	Leu	Val	Phe	Met	Ala	Ala	Ala	Pro	Ile	Leu	Ser	Ala	Phe	Arg	Leu	
		230					235					240				
												•				
ttc	tac	ttc	ggc	act	tac	ctg	cca	cac	aag	cct	gag	cca	ggc	cct	gca	944
Phe	Tyr	Phe	Gly	Thr	Tyr	Leu	Pro	His	Lys	Pro	Glu	Pro	Gly	Pro	Ala	
	245					250					255					
gca	ggc	tct	cag	gtg	atg	ġcc	tgg	ttc	agg	gcc	aag	aca	agt	gag _.	gca	992
Ala	Gly	Ser	Gln	Val	Met	Ala	Trp	Phe	Arg	Ala	Lys	Thr	Ser	Glu	Ala	
260					265					270					275	
tct	gat	gtg	atg	agt	ttc	ctg	aca	tgc	tac	cac	ttt	gac	ctg	cac	tgg	1040
Ser	qzA	Val	Met	Ser	Phe	Leu	Thr	Суз	Tyr	His	Phe	Asp	Leu	His	Trp	
				280					285					290		
gag	cac	cac	agg	tgg	ccc	ttt	gcc	ccc	tgg	tgg	cag	ctg	ccc	cac	tgc	1088
Glu	His	His	Arg	Trp	Pro	Phe	Ala	Pro	Trp	Trp	Gln	Leu	Pro	His	Cys	
			295					300					305			

WO 2004/063359 PCT/EP2004/000099 37/357

cgc cgc ctg	tcc ggg cgt g	gc ctg gtg cct	gcc ttg gca tg	1130
Arg Arg Leu	Ser Gly Arg G	ly Leu Val Pro	Ala Leu Ala	
310		315	320	

cctggtccct ccgctggtga cccagcgtct gcacaagagt gtcatgctac agggtgctgc 1190 ggccagtggc agcgcagtgc actctcagcc tgtatggggc taccgctgtg ccactgagca 1250 ' ctgggcatgc cactgagcac tgggcgtgct actgagcaat gggcgtgcta ctgagcaatg 1310 ggcgtgctac tgacaatggg cgtgctactg gggtctggca gtggctagga tggagtttga 1370 tgcattcagt agcggtggcc aacgtcatgt ggatggtgga agtgctgagg ggtttaggca 1430 gccggcattt gagagggcta agttataaat cgcatgctgc tcatgcgcac atatctgcac 1490 acagccaggg aaatcccttc gagagtgatt atgggacact tgtattggtt tcgtgctatt 1550 gttttattca gcagcagtac ttagtgaggg tgagagcagg gtggtgagag tggagtgagt 1610 gagtatgaac ctggtcagcg aggtgaacag cctgtaatga atgactctgt ct 1662

<210> 14

<211> 320

<212> PRT

<213> Haematococcus pluvialis

<400> 14

Met His Val Ala Ser Ala Leu Met Val Glu Gln Lys Gly Ser Glu Ala 1 5 10 15

Ala Ala Ser Ser Pro Asp Val Leu Arg Ala Trp Ala Thr Gln Tyr His
20 25 30

Met Pro Ser Glu Ser Ser Asp Ala Ala Arg Pro Ala Leu Lys His Ala 35 40 45

Tyr Lys Pro Pro Ala Ser Asp Ala Lys Gly Ile Thr Met Ala Leu Thr
50 55 60

Ile Ile Gly Thr Trp Thr Ala Val Phe Leu His Ala Ile Phe Gln Ile
65 70 75 80

Arg Leu Pro Thr Ser Met Asp Gln Leu His Trp Leu Pro Val Ser Glu 85 90 95

Ala Thr Ala Gln Leu Leu Gly Gly Ser Ser Ser Leu Leu His Ile Ala

100 105 110

Ala Val Phe Ile Val Leu Glu Phe Leu Tyr Thr Gly Leu Phe Ile Thr
115 120 125

Thr His Asp Ala Met His Gly Thr Ile Ala Leu Arg His Arg Gln Leu 130 135 140

Tyr Ser Met Leu His Arg Lys His Trp Glu His His Asn His Thr Gly
165 170 175

Glu Val Gly Lys Asp Pro Asp Phe His Lys Gly Asn Pro Gly Leu Val 180 185 190

Pro Trp Phe Ala Ser Phe Met Ser Ser Tyr Met Ser Leu Trp Gln Phe 195 200 205

Ala Arg Leu Ala Trp Trp Ala Val Val Met Gln Met Leu Gly Ala Pro

WO 2004/063359 PCT/EP2004/000099

39/357

210 215 220

Met Ala Asn Leu Leu Val Phe Met Ala Ala Pro Ile Leu Ser Ala 225 230 235 240

Phe Arg Leu Phe Tyr Phe Gly Thr Tyr Leu Pro His Lys Pro Glu Pro 245 250 255

Gly Pro Ala Ala Gly Ser Gln Val Met Ala Trp Phe Arg Ala Lys Thr 260 265 270

Ser Glu Ala Ser Asp Val Met Ser Phe Leu Thr Cys Tyr His Phe Asp 275 280 285

Leu His Trp Glu His His Arg Trp Pro Phe Ala Pro Trp Trp Gln Leu 290 295 300

Pro His Cys Arg Arg Leu Ser Gly Arg Gly Leu Val Pro Ala Leu Ala 305 310 315 320

<210> 15

<211> 729

<212> DNA

<213> Agrobacterium aurantiacum

<220>

<221> CDS

<222> (1)..(729)

<400> 15

atg agc gca cat gcc ctg ccc aag gca gat ctg acc gcc acc agc ctg

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

1 5 10 15

atc	gtc	tcg	ggc	ggc	atc	atc	gcc	gct	tgg	ctg	gcc	ctg	cat	gtg	cat	96
Ile	Val	Ser	Gly	Gly	Ile	Ile	Ala	Ala	Trp	Leu	Ala	Ėeu	His	Val	His	
	•		20					25		•			30			
gcg	ctg	tgg	ttt	ctg	gac	gça	gcg	gcg	cat	ccc	atc	cta	aca	atc	qca	144
	Leu															
		35			_		40					45				
	.1											43				
aat	ttc	cta	ααα	cta	acc	taa	ata	taa	ata	~~~	.					100
																192
ASII	Phe	пеп	GIŞ	Deu	1111		ьец	ser	vai	GIY		Pne	тте	тте	Ala	
	50					55					60					
	gac															240
His	Asp	Ala	Met	His	Gly	Ser	Val	Val	Pro	Gly	Arg	Pro	Arg	Ala	Asn	
65					70					75					80	
gcg	gcg	atg	ggc	cag	ctt	gtc	ctg	tgg	ctg	tat	gcc	gga	ttt	tcg	tgg	288
Ala	Ala	Met	Gly	Gln	Leu	Val	Leu	Trp	Leu	Tyr	Ala	Gly	Phe	Ser	Trp	
				85					90					95		
cgc	aag	atg	atc	atc	aaq	cac	ato	acc	cat	cac	cac	cat	acc	aaa	acc	336
	Lys												_			550
	-		100					105			9		110	CIJ	****	
			200					105					110			
~ 3.0	~~~	~~~		~~+	 -											
	gac															384
ASD	Asp		PIO	ASD	Pne	Asp		GTĀ	GIY	Pro	Val		Trp	Tyr	Ala	
		115					120					125				
	ttc															432
Arg	Phe	Ile	Gly	Thr	Tyr	Phe	Gly	Trp	Arg	Glu	Gly	Leu	Leu	Leu	Pro	
	130					135					140					
gtc	atc	gtg	acg	gtc	tat	gcg	ctg	atc	ctt	ggg	gat	cgc	tgg	atg	tac	480
Val	Ile	Val	Thr	Val	Tyr	Ala	Leu	Ile	Leu	Gly	Asp	Arg	Trp	Met	Tyr	
145					150					155					160	•
gtg	gtc	ttc	tgg	ccg	ctg	ccg	tcg	atc	ctg	gcg	tcg	atc	caq	cta	ttc	528
	Val															
			-	165				-	170					175		
									_, 5					.,J		
at.a	ttc	aac	acc	taa	cta	cca	cac	cac	ccc	aac	C2.C		~^~		000	E76
															_	576
VAL	Phe	GTĀ	T71T.	rrp	nen	PLO	nlS	wrg	rro	чтĀ	nlS	Asp	ALA	Lue	Pro	

WO 2004/063359		PCT/EP2004/000099
	41/357	

	W	2004	/0633	59						PCT/EP2004/0000						
									41/3	157						
			180					185					190			
gac	cgc	cac	aat	gcg	cgg	tcg	tcg	cgg	atc	agc	gac	ccc	gtg	tcg	ctg	624
Asp	Arg	His	Asn	Ala	Arg	Ser	Ser	Arg	Ile	Ser	Asp	Pro	Val	Ser	Leu	
		195					200					205				
ctg	acc	tgc	ttt	cac	ttt	ggc	ggt	tat	cat	cac	gaa	cac	cac	ctg	cac	672
Leu	Thr	Cys	Phe	His	Phe	Gly	Gly	Tyr	His	His	Glu	His	His	Leu	His	
	210					215					220					
ccg	acg	gtg	ccg	tgg	tgg	cgc	ctg	ccc	agc	acc	cgc	acc	aag	ggg	gac	720
Pro	Thr	Val	Pro	\mathtt{Trp}	\mathtt{Trp}	Arg	Leu	Pro	Ser	Thr	Arg	Thr	Lys	Gly	Asp	
225					230					235					240	
acc	gca	tga														729
Thr	Ala															
<210)>	1.6														
<211	>	242														
<212	> :	PRT														
<213	S> 2	Agrol	oacte	erium	aur	anti	acun	n								
<400)> :	16														

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Ile Ala

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Pro

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp

225 230 235 240

Thr Ala

<210> 17

<211> 1631

<212> DNA

<213> Alcaligenes sp.

<220>

<221> CDS

<222> (99)..(827)

<400> 17

ctgcaggccg ggcccggtgg ccaatggtcg caaccggcag gactggaaca ggacggcggg 60

ccggtctagg ctgtcgccct acgcagcagg agtttcgg atg tcc gga cgg aag cct 116

Met Ser Gly Arg Lys Pro

ı :

ggc aca act ggc gac acg atc gtc aat ctc ggt ctg acc gcc gcg atc 164
Gly Thr Thr Gly Asp Thr Ile Val Asn Leu Gly Leu Thr Ala Ala Ile

10 15 20

ctg ctg tgc tgg ctg gtc ctg cac gcc ttt acg cta tgg ttg cta gat 212 Leu Leu Cys Trp Leu Val Leu His Ala Phe Thr Leu Trp Leu Leu Asp

25 30 35

gcg gcc gcg cat ccg ctg ctt gcc gtg ctg tgc ctg gct ggg ctg acc 260

Ala Ala Ala His Pro Leu Leu Ala Val Leu Cys Leu Ala Gly Leu Thr

40 45 50

tgg ctg tcg gtc ggg ctg ttc atc atc gcg cat gac gca atg cac ggg 308
Trp Leu Ser Val Gly Leu Phe Ile Ile Ala His Asp Ala Met His Gly

55 60 65 70

tcc gtg gtg ccg ggg ccg ccg cgc gcc aat gcg gcg atc ggg caa ctg 356 Ser Val Val Pro Gly Arg Pro Arg Ala Asn Ala Ile Gly Gln Leu

WO 2004/063359		PCT/EP2004/000099
	4.4/2.55	

44/357 gcg ctg tgg ctc tat gcg ggg ttc tcg tgg ccc aag ctg atc gcc aag Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp Pro Lys Leu Ile Ala Lys cac atg acg cat cac cgg cac gcc ggc acc gac aac gat ccc gat ttc His Met Thr His His Arg His Ala Gly Thr Asp Asn Asp Pro Asp Phe ggt cac gga ggg ccc gtg cgc tgg tac ggc agc ttc gtc tcc acc tat Gly His Gly Gly Pro Val Arg Trp Tyr Gly Ser Phe Val Ser Thr Tyr ttc ggc tgg cga gag gga ctg ctg cta ccg gtg atc gtc acc acc tat Phe Gly Trp Arg Glu Gly Leu Leu Pro Val Ile Val Thr Thr Tyr gcg ctg atc ctg ggc gat cgc tgg atg tat gtc atc ttc tgg ccg gtc Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr Val Ile Phe Trp Pro Val ccg gcc gtt ctg gcg tcg atc cag att ttc gtc ttc gga act tgg ctg Pro Ala Val Leu Ala Ser Ile Gln Ile Phe Val Phe Gly Thr Trp Leu ccc cac cgc ccg gga cat gac gat ttt ccc gac cgg cac aac gcg agg Pro His Arg Pro Gly His Asp Asp Phe Pro Asp Arg His Asn Ala Arg tog acc ggc atc ggc gac cog ttg toa cta ctg acc tgc ttc cat ttc Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu Leu Thr Cys Phe His Phe

ggc ggc tat cac cac gaa cat cac ctg cat ccg cat gtg ccg tgg tgg

788

Gly Gly Tyr His His Glu His His Leu His Pro His Val Pro Trp

225

230

cgc ctg cct cgt aca cgc aag acc gga ggc cgc gca tga cgcaattcct 837
Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly Arg Ala

235 240

•			• •			
cattgtcgtg	gcgacagtcc	tcgtgatgga	gctgaccgcc	tattccgtcc	accgctggat	897
tatgcacggc	cccctaggct	ggggctggca	caagtcccat	cacgaagagc	acgaccacgc	957
gttggagaag	aacgacctct	acggcgtcgt	cttcgcggtg	ctggcgacga	tcctcttcac	1017
cgtgggcgcc	tattggtggc	cggtgctgtg	gtggatcgcc	ctgggcatga	cggtctatgg	1077
			·			
gttgatctat	ttcatcctgc	acgacgggct	tgtgcatcaa	cgctggccgt	ttcggtatat	1137
tccgcggcgg	ggctatttcc	gcaggctcta	ccaagctcat	cgcctgcacc	acgcggtcga	1197
ggggcgggac	cactgcgtca	gcttcggctt	catctatgcc	ccacccgtgg	acaagctgaa	1257
gcaggatctg	aagcggtcgg	gtgtcctgcg	ccccaggac	gagcgtccgt	cgtgatctct	1317
gateceggeg	tggccgcatg	aaatccgacg	tgctgctggc	aggggccggc	cttgccaacg	1377
gactgatcgc	gctggcgatc	cgcaaggcgc	ggcccgacct	tegegtgetg	ctgctggacc	1437
gtgcggcggg	cgcctcggac	gggcatactt	ggtcctgcca	cgacaccgat	ttggcgccgc	1497
actggctgga	ccgcctgaag	ccgatcaggc	gtggcgactg	gcccgatcag	gaggtgcggt	1557
tcccagacca	ttcgcgaagg	ctccgggccg	gatatggctc	gatcgacggg	cgggggctga	1617
tgcgtgcggt	gacc					1631

<210> 18

<211> 242

<212> PRT

<213> Alcaligenes sp.

<400> 18

Met Ser Gly Arg Lys Pro Gly Thr Thr Gly Asp Thr Ile Val Asn Leu 1 5 10 15

Gly Leu Thr Ala Ala Ile Leu Leu Cys Trp Leu Val Leu His Ala Phe

4076

20 25 30

Thr Leu Trp Leu Leu Asp Ala Ala Ala His Pro Leu Leu Ala Val Leu 35 40 45

Cys Leu Ala Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala
50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn 65 70 75 80

Ala Ala Ile Gly Gln Leu Ala Leu Trp Leu Tyr Ala Gly Phe Ser Trp 85 90 95

Pro Lys Leu Ile Ala Lys His Met Thr His His Arg His Ala Gly Thr 100 105 110

Asp Asn Asp Pro Asp Phe Gly His Gly Gly Pro Val Arg Trp Tyr Gly
115 120 125

Ser Phe Val Ser Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Leu Pro 130 135 140

Val Ile Val Thr Thr Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr
145 150 155 160

Val Ile Phe Trp Pro Val Pro Ala Val Leu Ala Ser Ile Gln Ile Phe 165 170 175

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Asp Phe Pro 180 185 190

Asp Arg His Asn Ala Arg Ser Thr Gly Ile Gly Asp Pro Leu Ser Leu 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220

Pro His Val Pro Trp Trp Arg Leu Pro Arg Thr Arg Lys Thr Gly Gly 225 230 235 240

Arg Ala

<210> 19

<211> 729

<212> DNA

<213> Paracoccus marcusii

<220>

<221> CDS

<222> (1)..(729)

<400> 19

atg age gea cat gee etg eee aag gea gat etg ace gee aca age etg 48 Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu 1 5 10 15

atc gtc tcg ggc ggc atc atc gcc gca tgg ctg gcc ctg cat gtg cat 96 Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His 20 25 30

gcg ctg tgg ttt ctg gac gcg gcc cat ccc atc ctg gcg gtc gcg 144 Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala 35 40 45

aat ttc ctg ggg ctg acc tgg ctg tcg gtc gga ttg ttc atc atc gcg 192 Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala

50 55 60

cat	gac	gcg	atg	cac	ggg	tcg	gtc	gtg	ccg	ggg	cgt	ccg	cac	acc	aat	240
															Asn	
65					70					75	3		3		80 .	
															00	
	gcg														-	288
Ala	Ala	Met	Gly	Gln	Leu	Val	Leu	Trp	Leu	Tyr	Ala	Gly	Phe	Ser	Trp	
				85					90					95		
cgc	aag	atg	atc	gtc	aag	cac	atg	gcc	cat	cac	cgc	cat	gcc	gga	acc	336
Arg	Lys	Met	Ile	Val	Lys	His	Met	Ala	His	His	Arg	His	Ala	Gly	Thr	
			100					105					110			
	gac															384
Asp	Asp	qaA	Pro	Asp	Phe	Asp	His	Gly	Gly	Pro	Val	Arg	Trp	Tyr	Ala	
		115					120					125				
											•					
	ttc															432
Arg	Phe	Ile	Gly	Thr	Tyr	Phe	Gly	Trp	Arg	Glu	Gly	Leu	Leu	Leu	Pro	
	130					135					140					
	atc															480
Va1	Ile	Val	Thr	Val	Tyr	Ala	Leu	Ile	Leu	Gly	Asp	Arg	Trp	Met	Tyr	
145					150					155					160	
	gtc															528
Val	Val	Phe	Trp	Pro	Leu	Pro	Ser	Ile	Leu	Ala	Ser	Ile	Gln	Leu	Phe	
				165					170					175		
	ttc															576
vaı	Phe	GIŸ		Trp	Leu	Pro	His		Pro	Gly	His	Asp	Ala	Phe	Pro	
			180					185					190			
a 20	966	a > t	-	~~~		L										
	cgc															624
АБР	Arg		ASII	Ala	Arg	ser		Arg	TTE	Ser	Asp		Val	Ser	Leu	
		195					200	•				205				
cta	acc	tac	ttt	cat	+++	aac	aa+	tat	ca+	C2C	~ = ^	C 2 C	a a.	a+~	a na	670
	Thr															672
سان س	210	-10		-440		215	GTĀ	TÄT	1172	uts	220	nis	uis	neu	nis	
						443					22U					

WO 2004/063359 PCT/EP2004/000099

49/357

ccg acg gtg ccg tgg tgg cgc ctg ccc agc acc cgc acc aag ggg gac 720
Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp

225 230 235 240

acc gca tga 729

Thr Ala

<210> 20

<211> 242

<212> PRT

<213> Paracoccus marcusii

<400> 20

Met Ser Ala His Ala Leu Pro Lys Ala Asp Leu Thr Ala Thr Ser Leu

1 5 10 15

Ile Val Ser Gly Gly Ile Ile Ala Ala Trp Leu Ala Leu His Val His
20 25 30

Ala Leu Trp Phe Leu Asp Ala Ala Ala His Pro Ile Leu Ala Val Ala 35 40 45

Asn Phe Leu Gly Leu Thr Trp Leu Ser Val Gly Leu Phe Ile Ile Ala 50 55 60

His Asp Ala Met His Gly Ser Val Val Pro Gly Arg Pro Arg Ala Asn 65 70 75 80

Ala Ala Met Gly Gln Leu Val Leu Trp Leu Tyr Ala Gly Phe Ser Trp

85 90 95

Arg Lys Met Ile Val Lys His Met Ala His His Arg His Ala Gly Thr
100 105 110

Asp Asp Asp Pro Asp Phe Asp His Gly Gly Pro Val Arg Trp Tyr Ala 115 120 125

Arg Phe Ile Gly Thr Tyr Phe Gly Trp Arg Glu Gly Leu Leu Pro 130 135 140

Val Ile Val Thr Val Tyr Ala Leu Ile Leu Gly Asp Arg Trp Met Tyr 145 150 155 160

Val Val Phe Trp Pro Leu Pro Ser Ile Leu Ala Ser Ile Gln Leu Phe 165 170 ·

Val Phe Gly Thr Trp Leu Pro His Arg Pro Gly His Asp Ala Phe Pro 180 185 190

Asp Arg His Asn Ala Arg Ser Ser Arg Ile Ser Asp Pro Val Ser Leu 195 200 205

Leu Thr Cys Phe His Phe Gly Gly Tyr His His Glu His His Leu His 210 215 220

Pro Thr Val Pro Trp Trp Arg Leu Pro Ser Thr Arg Thr Lys Gly Asp 225 230 235 240

Thr Ala

<210> 21

<211> 1629

<212> DNA

<213> Synechocystis sp.

<220> <221> CDS <222> (1)..(1629)<400> 21 atg atc acc acc gat gtt gtc att att ggg gcg ggg cac aat ggc tta 48 Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu 10 15 gtc tgt gca gcc tat ttg ctc caa cgg ggc ttg ggg gtg acg tta cta 96 Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu 20 25 30 gaa aag cgg gaa gta cca ggg ggg gcg gcc acc aca gaa gct ctc atg 144 Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met 35 40 45 ccg gag cta tcc ccc cag ttt cgc ttt aac cgc tgt gcc att gac cac 192 Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His 50 55 60 gaa ttt atc ttt ctg ggg ccg gtg ttg cag gag cta aat tta gcc cag 240 Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln 70 65 75 80 tat ggt ttg gaa tat tta ttt tgt gac ccc agt gtt ttt tgt ccg ggg 288 Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly 85 90 95 ctg gat ggc caa gct ttt atg agc tac cgt tcc cta gaa aaa acc tgt 336 Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys 100 105 110 gcc cac att gcc acc tat agc ccc cga gat gcg gaa aaa tat cgg caa 384 Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln 115 120 125 ttt gtc aat tat tgg acg gat ttg ctc aac gct gtc cag cct gct ttt 432

Phe	Val		Tyr	Trp	Thr	Asp 135		Leu	Asr	a Ala	Val		Pro	Ala	Phe .	
Asn	Ala				Ala	Leu				Ala	Leu				tgg Trp	480
145 gaa		tta	aaa	tcc	150 gtg		gcg	atc	gcc	155 ggg		aaa	acc	aag	160 gcg	528
					Val					Gly					Ala	
												gat Asp				576
												gct Ala 205	_		_	624
												agc Ser			_	672
												aga Arg				720
												gtg Val				768
Gly												gta Val				816
aac Asn	Asn					Val										864
gcc Ala															_	912

300

53/357 295

.290

450

455

460

caa	ttg	gtg	gaa	ccg	ggg	gcc	cta	gcc	aag	gtg	aat	caa	aac	cta	ggg	960
Gln	Leu	Val	Glu	Pro	Gly	Ala	Leu	Ala	Lys	Val	Asn	Gln	Asn	Leu	${ t Gly}$	
305					310					315					320	
								•								
gaa	cga	ctg	gaa	cgg	cgc	act	gtg	aac	aat	aac	gaa	gcc	att	tta	aaa	1008
Glu	Arg	Leu	Glu	Arg	Arg	Thr	Val	Asn	Asn	Asn	Glu	Ala	Ile	Leu	Lys	
				325	•				330					335		
	•															
atc	gat	tgt	gcc	ctc	tcc	ggt	tta	ccc	cac	ttc	act	gcc	atg	gcc	ggg	1056
Ile	Asp	Cys	Ala	Leu	Ser	Gly	Leu	Pro	His	Phe	Thr	Ala	Met	Ala	Gly	
			340					345					350			
ccg	gag	gat	cta	acg	gga	act	att	ttg	att	gcc	gac	tcg	gta	cgc	cat	1104
Pro	Glu	Asp	Leu	Thr	Gly	Thr	Ile	Leu	Ile	Ala	Asp	Ser	Val	Arg	His	
		355					360					365				
gtc	gag	gaa	gcc	cac	gcc	ctc	att	gcc	ttg	ggg	caa	att	ccc	gat	gct	1152
Val	Glu	Glu	Ala	His	Ala	Leu	Ile	Ala	Leu	Gly	Gln	Ile	Pro	Asp	Ala	
	370					375					380					
aat	ccg	tct	tta	tat	ttg	gat	att	ccc	act	gta	ttg	gac	ccc	acc	atg	1200
Asn	Pro	Ser	Leu	Tyr	Leu	Asp	Ile	Pro	Thr	Val	Leu	Asp	Pro	Thr	Met	
385					390					395					400	
gcc	ccc	cct	ggg	cag	cac	acc	ctc	tgg	atc	gaa	ttt	ttt	gcc	ccc	tac	1248
Ala	Pro	Pro	Gly	Gln	His	Thr	Leu	Trp	Ile	Glu	Phe	Phe	Ala	Pro	Tyr	
				405					410					415		
cgc	atc	gcc	ggg	ttg	gaa	ggg	aca	ggg	tta	atg	ggc	aca	ggt	tgg	acc	1296
Arg	Ile	Ala	Gly	Leu	Glu	Gly	Thr	Gly	Leu	Met	Gly	Thr	Gly	Trp	Thr	
			420					425					430			
•																
gat	gag	tta	aag	gaa	aaa	gtg	gcg	gat	cgg	gtg	att	gat	aaa	tta	acg	1344
Asp	Glu	Leu	Lys	Glu	Lys	Val	Ala	Asp	Arg	Val	Ile	Asp	Lys	Leu	Thr	
		435					440					445				
gac	tat	gcc	cct	aac	cta	aaa	tct	ctg	atc	att	ggt	cgc	cga	gtg	gaa	1392
Asp	Tyr	Ala	Pro	Asn	Leu	Lys	Ser	Leu	Ile	Ile	Gly	Arg	Arg	Val	G1u	

- 1,	
	•

agt ccc gcc gaa ctg gcc caa cgg ctg gga agt tac aac ggc aat gtc 1440 Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val 465

tat cat ctg gat atg agt ttg gac caa atg atg ttc ctc cgg cct cta 1488

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu

485

490

495

ccg gaa att gcc aac tac caa acc ccc atc aaa aat ctt tac tta aca 1536

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr
500 505 510

ggg gcg ggt acc cat ccc ggt ggc tcc ata tca ggt atg ccc ggt aga 1584
Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg .

515 520 525

aat tgc gct cgg gtc ttt tta aaa caa cgt cgt ttt tgg taa 1629
Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp
530 540

<210> 22

<211> 542

<212> PRT

<213> Synechocystis sp.

<400> 22

Met Ile Thr Thr Asp Val Val Ile Ile Gly Ala Gly His Asn Gly Leu
1 5 10 15

Val Cys Ala Ala Tyr Leu Leu Gln Arg Gly Leu Gly Val Thr Leu Leu 20 25 30

Glu Lys Arg Glu Val Pro Gly Gly Ala Ala Thr Thr Glu Ala Leu Met

35 40 45

Pro Glu Leu Ser Pro Gln Phe Arg Phe Asn Arg Cys Ala Ile Asp His

Glu Phe Ile Phe Leu Gly Pro Val Leu Gln Glu Leu Asn Leu Ala Gln

Tyr Gly Leu Glu Tyr Leu Phe Cys Asp Pro Ser Val Phe Cys Pro Gly 90 -

Leu Asp Gly Gln Ala Phe Met Ser Tyr Arg Ser Leu Glu Lys Thr Cys

Ala His Ile Ala Thr Tyr Ser Pro Arg Asp Ala Glu Lys Tyr Arg Gln

Phe Val Asn Tyr Trp Thr Asp Leu Leu Asn Ala Val Gln Pro Ala Phe

Asn Ala Pro Pro Gln Ala Leu Leu Asp Leu Ala Leu Asn Tyr Gly Trp

Glu Asn Leu Lys Ser Val Leu Ala Ile Ala Gly Ser Lys Thr Lys Ala

Leu Asp Phe Ile Arg Thr Met Ile Gly Ser Pro Glu Asp Val Leu Asn

Glu Trp Phe Asp Ser Glu Arg Val Lys Ala Pro Leu Ala Arg Leu Cys

Ser Glu Ile Gly Ala Pro Pro Ser Gln Lys Gly Ser Ser Ser Gly Met

220

30

215

210

Met Met Val Ala Met Arg His Leu Glu Gly Ile Ala Arg Pro Lys Gly 225 230 235 240

Gly Thr Gly Ala Leu Thr Glu Ala Leu Val Lys Leu Val Gln Ala Gln 245 250 255

Gly Gly Lys Ile Leu Thr Asp Gln Thr Val Lys Arg Val Leu Val Glu 260 265 270

Asn Asn Gln Ala Ile Gly Val Glu Val Ala Asn Gly Glu Gln Tyr Arg 275 280 285

Ala Lys Lys Gly Val Ile Ser Asn Ile Asp Ala Arg Arg Leu Phe Leu 290 295 300

Gln Leu Val Glu Pro Gly Ala Leu Ala Lys Val Asn Gln Asn Leu Gly 305 310 315 320

Glu Arg Leu Glu Arg Arg Thr Val Asn Asn Glu Ala Ile Leu Lys
. 325 330 335

Ile Asp Cys Ala Leu Ser Gly Leu Pro His Phe Thr Ala Met Ala Gly 340 345 350

Pro Glu Asp Leu Thr Gly Thr Ile Leu Ile Ala Asp Ser Val Arg His 355 360 365

Val Glu Glu Ala His Ala Leu Ile Ala Leu Gly Gln Ile Pro Asp Ala 370 375 380

Asn Pro Ser Leu Tyr Leu Asp Ile Pro Thr Val Leu Asp Pro Thr Met

Ala Pro Pro Gly Gln His Thr Leu Trp Ile Glu Phe Phe Ala Pro Tyr

Arg Ile Ala Gly Leu Glu Gly Thr Gly Leu Met Gly Thr Gly Trp Thr

Asp Glu Leu Lys Glu Lys Val Ala Asp Arg Val Ile Asp Lys Leu Thr

Asp Tyr Ala Pro Asn Leu Lys Ser Leu Ile Ile Gly Arg Arg Val Glu

Ser Pro Ala Glu Leu Ala Gln Arg Leu Gly Ser Tyr Asn Gly Asn Val

Tyr His Leu Asp Met Ser Leu Asp Gln Met Met Phe Leu Arg Pro Leu

Pro Glu Ile Ala Asn Tyr Gln Thr Pro Ile Lys Asn Leu Tyr Leu Thr . 510

Gly Ala Gly Thr His Pro Gly Gly Ser Ile Ser Gly Met Pro Gly Arg

Asn Cys Ala Arg Val Phe Leu Lys Gln Gln Arg Arg Phe Trp

<211> 776

<212> DNA

<213> Bradyrhizobium sp.

<220>

<221> CDS

<222> (1)..(774)

<400> 23

atg cat gca gca acc gcc aag gct act gag ttc ggg gcc tct cgg cgc 48

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg

1 5 10 15

gac gat gcg agg cag cgc cgc gtc ggt ctc acg ctg gcc gcg gtc atc 96
Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile
20 25 30

atc gcc gcc tgg ctg gtg ctg cat gtc ggt ctg atg ttc ttc tgg ccg

144

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro

35

40

45

acc tgg ctc tat gta ggc ctg ttc atc atc gcg cat gac tgc atg cac

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His

70

75

80

ggc tcg ctg gtg ccg ttc aag ccg cag gtc aac cgc cgt atc gga cag 288
Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln
85 90 95

ctc tgc ctg ttc ctc tat gcc ggg ttc tcc ttc gac gct ctc aat gtc 336
Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val
100 105 110

gag cac cac aag cat cac cgc cat ccc ggc acg gcc gag gat ccc gat

384
Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp

115
120
125

59/357

٠.									02,	• • •						
ttc	gac	gag	gtg	ccg	ccg	cac	ggc	ttc	tgg	cac	tgg	ttc	gcc	agc	ttt	432
Phe	Asp	Glu	Val	Pro	Pro	His	Gly	Phe	Trp	His	Trp	Phe	Ala	Ser	Phe	
	130					135					140					
ttc	ctg	cac	tat	ttc	ggc	tgg	aag	cag	gtc	gcg	atc	atc	gca	gcc	gtc	480
Phe	Leu	His	Tyr	Phe	Gly	Trp	Lys	Gln	Val	Ala	Ile	Ile	Ala	Ala	Va1	
145					150					155					160	
tcg	ctg	gtt	tat	cag	ctc	gtc	ttc	gcc	gtt	ccc	ttg	cag	aac	atc	ctg	528
Ser	Leu	Val	Tyr	Gln	Leu	Val	Phe	Ala	Val	Pro	Leu	Gln	Asn	Ile	Leu	
				165					170		•			175	•	
ctg	ttc	tgg	gcg	ctg	ccc	ggg	ctg	ctg	tcg	gcg	ctg	cag	ctg	ttc	acc	576
Leu	Phe	Trp	Ala	Leu	Pro	Gly	Leu	Leu	Ser	Ala	Leu	Gln	Leu	Phe	Thr	
			180					185					190			
ttc	ggc	acc	tat	ctg	ccg	cac	aag	ccg	gcc	acg	cag	CCC	ttc	gcc	gat	624
Phe	Gly	Thr	Tyr	Leu	Pro	His	Lys	Pro	Ala	Thr	Gln	Pro	Phe	Ala	Asp	
		195					200	•				205				
	•															
cgc	cac	aac	gcg	cgg	acg	agc	gaa	ttt	ccc	gcg	tgg	ctg	tcg	ctg	ctg	672
Arg	His	Asn	Ala	Arg	Thr	Ser	Glu	Phe	Pro	Ala	Trp	Leu	Ser	Leu	Leu	
	210					215					220					
acc	tgc	ttc	cac	ttc	ggc	ttt	cat	cac	gag	cat	cat	ctg	cat	ccc	gat	720
Thr	Cys	Phe	His	Phe	Gly	Phe	His	His	Glu	His	His	Leu	His	Pro	Asp	
225					230					235					240	
gcg	ccg	tgg	tgg	cgg	ctg	ccg	gag	atc	aag	cgg	cgg	gcc	ctg	gaa	agg	768
Ala	Pro	Trp	Trp	Arg	Leu	Pro	Glu	Ile	Lys	Arg	Arg	Ala	Leu	Glu	Arg	
				245					250					255		
cgt	gac	ta														776
Arg	qaA															

<210> 24

<211> 258

<212> PRT

<213> Bradyrhizobium sp.

<400> 24

Met His Ala Ala Thr Ala Lys Ala Thr Glu Phe Gly Ala Ser Arg Arg

1 10 15

Asp Asp Ala Arg Gln Arg Arg Val Gly Leu Thr Leu Ala Ala Val Ile 20 25 30

Ile Ala Ala Trp Leu Val Leu His Val Gly Leu Met Phe Phe Trp Pro 35 40 45

Leu Thr Leu His Ser Leu Leu Pro Ala Leu Pro Leu Val Val Leu Gln 50 55 60

Thr Trp Leu Tyr Val Gly Leu Phe Ile Ile Ala His Asp Cys Met His 65 70 75 80

Gly Ser Leu Val Pro Phe Lys Pro Gln Val Asn Arg Arg Ile Gly Gln 85 90 95

Leu Cys Leu Phe Leu Tyr Ala Gly Phe Ser Phe Asp Ala Leu Asn Val 100 105 110

Glu His His Lys His His Arg His Pro Gly Thr Ala Glu Asp Pro Asp 115 120 125

Phe Asp Glu Val Pro Pro His Gly Phe Trp His Trp Phe Ala Ser Phe 130 135 140

 WO 2004/063359 PCT/EP2004/000099

61/357

Ser Leu Val Tyr Gln Leu Val Phe Ala Val Pro Leu Gln Asn Ile Leu 165 170 175

Leu Phe Trp Ala Leu Pro Gly Leu Leu Ser Ala Leu Gln Leu Phe Thr
180 185 190

Phe Gly Thr Tyr Leu Pro His Lys Pro Ala Thr Gln Pro Phe Ala Asp 195 200 205

Arg His Asn Ala Arg Thr Ser Glu Phe Pro Ala Trp Lèu Ser Leu Leu 210 215 220

Thr Cys Phe His Phe Gly Phe His His Glu His His Leu His Pro Asp 225 230 · 235 240

Ala Pro Trp Trp Arg Leu Pro Glu Ile Lys Arg Arg Ala Leu Glu Arg
245 250 255

Arg Asp

<210> 25

<211> 777

<212> DNA

<213> Nostoc sp.

<220>

<221> CDS

<222> (1)..(777)

<400> 25

atg gtt cag tgt caa cca tca tct ctg cat tca gaa aaa ctg gtg tta Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu

1 5 10 15

ttg	tca	tcg	aca	atc	aga	gat	gat	aaa	aat	att	aat	aag	ggt	ata	ttt	96
Leu	Ser	Ser	Thr	Ile	Arg	Asp	Asp	Lys	Asn	Ile	Asn	Lys	Gly	Ile	Phe	
			20					25		•			30			
att	gcc	tgc	ttt	atc	tta	ttt	tta	tgg	gca	att	agt	tta	atc	tta	tta	144
															Leu	
		35					40	-				45				
ctc	tca	atà	gat	aca	tcc	ata	att	cat	220	200	++=	tta	aat	ata	acc	192
		Ile														192
Бец	50	116	nsp	1111	Ser		116	птѕ	цуъ	Ser		пеа	GTA	TTG	Ala	
	50					55					60					
		tgg												_		240
Met	Leu	Trp	Gln	Thr	Phe	Leu	Tyr	Thr	Gly	Leu	Phe	Ile	Thr	Ala	His	
65					70					75					80	
			•													
gat	gcc	atg	cac	ggc	gta	gtt	tat	ccc	aaa	aat	ccc	aga	ata	aat	aat	288
Asp	Ala	Met	His	Gly	Val	Val	Tyr	Pro	Lys	Asn	Pro	Arg	Ile	Asn	Asn	
				85					90					95		
ttt	ata	ggt	aag	ctc	act	cta	atc	ttg	tat	gga	cta	ctc	cct	tat	aaa	336
Phe	Ile	Gly	Lys	Leu	Thr	Leu	Ile	Leu	Tyr	Gly	Leu	Leu	Pro	Tyr	Lys	
			100					105					110			
gat	tta	ttg	aaa	aaa	cat	tgg	tta	cac	cac	gga	cat	cct	gat	act	gat	384
		Leu													_	
		115	-	_		_	120			_		125				
tta	gac	cct	aat	tat	tac	aat	aat	cat	ccc	caa	820	tta	+++	at t	taa	432
		Pro														432
Dea	130	FLO	Asp	TAT	TAT	135	GTĀ	птр	PIO	GIII		Pne	Pne	Leu	Trp	
	130					733					140					
			-	_ 4		1 k	L_L								•	
		cat														480
	ren	His	Pne	Met		ser	'l'Yr	Trp	Arg		Thr	Gln	Ile	Phe	Gly	
145					150					155					160	
		atg													_	528
Leu	Val	Met	Ile	Phe	His	Gly	Leu	Lys	Asn	Leu	Val	His	Ile	Pro	Glu	
				165					170					175		
aat	aat	tta	att	ata	ttt	tgg	atg	ata	cct	tct	att	tta	agt	tca	gta	576

WO 2004/063359 PCT/EP2004/000099 63/357

Asn	Asn	Leu	Ile	Ile	Phe	Trp	Met	Ile	Pro	Ser	Ile	Leu	Ser	Ser	Val
			180					185					190		

caa cta ttt tat ttt ggt aca ttt ttg cct cat aaa aag cta gaa ggt 624

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly
195 200 205

ggt tat act aac ccc cat tgt gcg cgc agt atc cca tta cct ctt ttt 672

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe

210 215 220

tgg tct ttt gtt act tgt tat cac ttc ggc tac cac aag gaa cat cac 720

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His

230 235 240

gaa tac cct caa ctt cct tgg tgg aaa tta cct gaa gct cac aaa ata 768
Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile
245 250 255

tct tta taa 777
Ser Leu

<210> 26 ·

<211> 258

<212> PRT

<213> Nostoc sp.

<400> 26

Met Val Gln Cys Gln Pro Ser Ser Leu His Ser Glu Lys Leu Val Leu 1 5 10 15

Leu Ser Ser Thr Ile Arg Asp Asp Lys Asn Ile Asn Lys Gly Ile Phe 20 25 30

Ile Ala Cys Phe Ile Leu Phe Leu Trp Ala Ile Ser Leu Ile Leu Leu
35 40 45

Leu Ser Ile Asp Thr Ser Ile Ile His Lys Ser Leu Leu Gly Ile Ala

Met Leu Trp Gln Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His

Asp Ala Met His Gly Val Val Tyr Pro Lys Asn Pro Arg Ile Asn Asn

Phe Ile Gly Lys Leu Thr Leu Ile Leu Tyr Gly Leu Leu Pro Tyr Lys

Asp Leu Leu Lys Lys His Trp Leu His His Gly His Pro Gly Thr Asp

Leu Asp Pro Asp Tyr Tyr Asn Gly His Pro Gln Asn Phe Phe Leu Trp

Tyr Leu His Phe Met Lys Ser Tyr Trp Arg Trp Thr Gln Ile Phe Gly

Leu Val Met Ile Phe His Gly Leu Lys Asn Leu Val His Ile Pro Glu

Asn Asn Leu Ile Ile Phe Trp Met Ile Pro Ser Ile Leu Ser Ser Val

Gln Leu Phe Tyr Phe Gly Thr Phe Leu Pro His Lys Lys Leu Glu Gly

WO 2004/063359 PCT/EP2004/000099 65/357

Gly Tyr Thr Asn Pro His Cys Ala Arg Ser Ile Pro Leu Pro Leu Phe 210 215 220

Trp Ser Phe Val Thr Cys Tyr His Phe Gly Tyr His Lys Glu His His 225 230 235 240

Glu Tyr Pro Gln Leu Pro Trp Trp Lys Leu Pro Glu Ala His Lys Ile 245 250 255

Ser Leu

<210> 27

<211> 789

<212> DNA

<213> Nostoc punctiforme

<220>

<221> CDS

<222> (1)..(789)

<400> 27

ttg aat ttt tgt gat aaa cca gtt agc tat tat gtt gca ata gag caa 48 Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln 1 5 10 15

tta agt gct aaa gaa gat act gtt tgg ggg ctg gtg att gtc ata gta 96
Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val
20 . 25 30

att att agt ctt tgg gta gct agt ttg gct ttt tta cta gct att aat 144

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn

35 40 45

tat gcc aaa gtc cca att tgg ttg ata cct att gca ata gtt tgg caa 192

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln

50 55 60

															cat	240
	: Phe	Leu	Тут	Thr	Gly	Leu	Phe	Ile	Thr	Ala	His	a Asp	Ala	Met	His	
65					70					75					80	
																•
															tca	288
Gly	Ser	Val	Tyr		Lys	Asn	Pro	Lys	Ile	Asn	Asn	Phe	Ile	Gly	Ser	
				85				•	90					95		
cta	act	at a	~~~	a++	+			1.1.				•				
		gta Val														336
Dea	*****	var	100	neu	TAT	AIG	val	105	PIO	TYL	GIN	GIN		ьeu	гÃЗ	
			100					103					110			
aat	cat	tgc	tta	cat	cat	cgt	cat	cct	gct	agc	gaa	att	gac	cca	gat.	384
		Cys														
		115					120					125	_		,	
ttt	cat	gat	ggt	aag	aga	aca	aac	gct	att	ttc	tgg	tat	ctc	cat	ttc	432
Phe	His	Asp	Gly	Lys	Arg	Thr	Asn	Ala	Ile	Phe	Trp	Tyr	Leu	His	Phe	
	130					135					140					
		gaa														480
	Ile	Glu	Tyr	Ser	Ser	Trp	Gln	Gln	Leu	Ile	Val	Leu	Thr	Ile	Leu	
145					150					155					160	
+++	22+	++-	~~+		.		L.L		,							
		tta														528
2210	11.511	Leu	AIG	165	TAT	vaı	ьeu	urs	170	HIS	GIN	тте	Asn		IIe	
				103					170					175	•	
tta	ttt	tgg	agt	att	cct	cca	att	tta	agt	tcc	att	caa	cta	ttt	tat	576
		Trp														3,0
			180					185					190		-4	
ttc	gga	aca	ttt	ttg	cct	cat	cga	gaa	ccc	aag	aaa	gga	tat	gtt	tat	624
Phe	Gly	Thr	Phe	Leu	Pro	His	Arg	Glu	Pro	Lys	Lys	Gly	Tyr	Val	Tyr	
		195					200					205				
		tgc														672
Pro		Суѕ	Ser	Gln			Lys	Leu	Pro			Leu	Ser	Phe	Ile	
	210					215					220					
act	tac	tac	cac	+++	aa+	+=+	aa+	722	~ ~ ~	~ ~+	an+	~~~	-			E00
gct	Lyc	tac	Cat	LLL	ggt	Lat	cat	yaa	yaa	cat	cat	yag	cat	ccc	cat	720

WO 2004/063359 PCT/EP2004/000099 67/357

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His

225 230 235 240

gta cct tgg tgg caa ctt cca tct gta tat aag cag aga gta ttc aac 768 Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn 245 250

aat tca gta acc aat tcg taa 789 Asn Ser Val Thr Asn Ser 260

<210> 28

<211> 262

<212> PRT

<213> Nostoc punctiforme

<400> 28

Leu Asn Phe Cys Asp Lys Pro Val Ser Tyr Tyr Val Ala Ile Glu Gln 5 10 15

Leu Ser Ala Lys Glu Asp Thr Val Trp Gly Leu Val Ile Val Ile Val . 30 20 25

Ile Ile Ser Leu Trp Val Ala Ser Leu Ala Phe Leu Leu Ala Ile Asn 35 40 45

Tyr Ala Lys Val Pro Ile Trp Leu Ile Pro Ile Ala Ile Val Trp Gln 50 55

Met Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ala His Asp Ala Met His 70 75 80

Gly Ser Val Tyr Arg Lys Asn Pro Lys Ile Asn Asn Phe Ile Gly Ser 90 85 95

Leu Ala Val Ala Leu Tyr Ala Val Phe Pro Tyr Gln Gln Met Leu Lys

Asn His Cys Leu His His Arg His Pro Ala Ser Glu Val Asp Pro Asp

Phe His Asp Gly Lys Arg Thr Asn Ala Ile Phe Trp Tyr Leu His Phe

Met Ile Glu Tyr Ser Ser Trp Gln Gln Leu Ile Val Leu Thr Ile Leu

Phe Asn Leu Ala Lys Tyr Val Leu His Ile His Gln Ile Asn Leu Ile 165 170

Leu Phe Trp Ser Ile Pro Pro Ile Leu Ser Ser Ile Gln Leu Phe Tyr

Phe Gly Thr Phe Leu Pro His Arg Glu Pro Lys Lys Gly Tyr Val Tyr

Pro His Cys Ser Gln Thr Ile Lys Leu Pro Thr Phe Leu Ser Phe Ile

Ala Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His

Val Pro Trp Trp Gln Leu Pro Ser Val Tyr Lys Gln Arg Val Phe Asn

260

<210> 29 <211> 762 <212> DNA <213> Nostoc punctiforme <220> <221> CDS <222> (1)..(762) <400> 29 gtg atc cag tta gaa caa cca ctc agt cat caa gca aaa ctg act cca 48 Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro 5 10 gta ctg aga agt aaa tct cag ttt aag ggg ctt ttc att gct att gtc 96 Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val 20 25 30 att gtt agc gca tgg gtc att agc ctg agt tta tta ctt tcc ctt gac 144 Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp 35 40 45 atc tca aag cta aaa ttt tgg atg tta ttg cct gtt ata cta tgg caa 192 Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln 50 55 60 aca ttt tta tat acg gga tta ttt att aca tct cat gat gcc atg cat 240 Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His 65 70 75 ggc gta gta ttt ccc caa aac acc aag att aat cat ttg att gga aca 288 Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr 90 85 ttg acc cta tcc ctt tat ggt ctt tta cca tat caa aaa cta ttg aaa 336 Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Leu Lys 100 105 110

70/357

aaa	cat	tgg	tta	cac	cac	cac	aat	cca	gca	agc	tca	ata	gac	ccg	gat	384
Lys	His	Trp	Leu	His	His	His	Asn	Pro	Ala	Ser	Ser	Ile	Asp	Pro	Asp	
		115					120					125				
			٠													
ttt	cac	aat	ggt	aaa	cac	caa	agt	ttc	ttt	gct	tgg	tat	ttt	cat	ttt	432
Phe	His	Asn	Gly	Lys	His	Gln	Ser	Phe	Phe	Ala	Trp	Tyr	Phe	His	Phe	
	130					135					140					
atg	aaa	ggt	tac	tgg	agt.	tgg	ggg	caa	ata	att	gcg	ttg	act	att	att	480
Met	Lvs	Glv	Tvr	Trp	Ser	Ψrp	Glv	Gln	al T	Tle	Δla	T.011	Thr	Tla	Tle	

atg aaa ggt tac tgg agt tgg ggg caa ata att gcg ttg act att att

480

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile

145

150

150

160

tat aac ttt gct aaa tac ata ctc cat atc cca agt gat aat cta act 528

Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr

165 170 175

tac ttt tgg gtg cta ccc tcg ctt tta agt tca tta caa tta ttc tat 576

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr

180 185 190 .

ttt ggt act ttt tta ccc cat agt gaa cca ata ggg ggt tat gtt cag

624

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln

195

200

205

cct cat tgt gcc caa aca att agc cgt cct att tgg tgg tca tft atc 672

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile
210 215 220

acg tgc tat cat ttt ggc tac cac gag gaa cat cac gaa tat cct cat

720

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His

225

230

240

att tct tgg tgg cag tta cca gaa att tac aaa gca aaa tag 762

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys
245 250

<210> 30

<211> 253

<212> PRT

<213> Nostoc punctiforme

71/357

<400> 30

Val Ile Gln Leu Glu Gln Pro Leu Ser His Gln Ala Lys Leu Thr Pro

Val Leu Arg Ser Lys Ser Gln Phe Lys Gly Leu Phe Ile Ala Ile Val

Ile Val Ser Ala Trp Val Ile Ser Leu Ser Leu Leu Ser Leu Asp

Ile Ser Lys Leu Lys Phe Trp Met Leu Leu Pro Val Ile Leu Trp Gln

Thr Phe Leu Tyr Thr Gly Leu Phe Ile Thr Ser His Asp Ala Met His

Gly Val Val Phe Pro Gln Asn Thr Lys Ile Asn His Leu Ile Gly Thr

Leu Thr Leu Ser Leu Tyr Gly Leu Leu Pro Tyr Gln Lys Leu Lys

Lys His Trp Leu His His Asn Pro Ala Ser Ser Ile Asp Pro Asp

Phe His Asn Gly Lys His Gln Ser Phe Phe Ala Trp Tyr Phe His Phe

Met Lys Gly Tyr Trp Ser Trp Gly Gln Ile Ile Ala Leu Thr Ile Ile Tyr Asn Phe Ala Lys Tyr Ile Leu His Ile Pro Ser Asp Asn Leu Thr

165 170 175

Tyr Phe Trp Val Leu Pro Ser Leu Leu Ser Ser Leu Gln Leu Phe Tyr
180 185 190

Phe Gly Thr Phe Leu Pro His Ser Glu Pro Ile Gly Gly Tyr Val Gln
195 200 205

Pro His Cys Ala Gln Thr Ile Ser Arg Pro Ile Trp Trp Ser Phe Ile 210 215 220

Thr Cys Tyr His Phe Gly Tyr His Glu Glu His His Glu Tyr Pro His 225 230 235 240

Ile Ser Trp Trp Gln Leu Pro Glu Ile Tyr Lys Ala Lys $245 \hspace{1.5cm} 250 \hspace{1.5cm}$

<210> 31

<211> 1608

<212> DNA

<213> Haematococcus pluvialis

<220>

<221> CDS

<222> (3)..(971)

<400> 31

ct aca ttt cac aag ccc gtg agc ggt gca agc gct ctg ccc cac atc

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile

1 5 10 15

95

ggc cca cct cct cat ctc cat cgg tca ttt gct gct acc acg atg ctg

73/357

Gly	Pro	Pro	Pro	His 20	Leu	His	Arg	Ser	Phe 25	Ala	Ala	Thr	Thr	Met 30	Leu	
tcg	aag	ctg	cag	tca	atc	agc	gtc	aag	gcc	cgc	cgc	gtt	gaa	cta	gcc	143
Ser	Lys	Leu	Gln	Ser	Ile	Ser	Val	Lys	Ala	Arg	Arg	Val	Glu	Leu	Ala	
			35					40					45			
cgc	gac	atc	acg	cgg	ccc	aaa	gtc	tgc	ctg	cat	gct	cag	cgg	tgc	tcg	191
Arg	Asp	Iḷe	Thr	Arg	Pro	Lys	Val	Cys	Leu	His	Ala	Gln	Arg	Cys	Ser	
		50					55					60				
tta	gtt	cgg	ctg	cga	gtg	gca	gca	cca	cag	aca	gag	gag	gcg	ctg	gga	239
Leu	Val	Arg	Leu	Arg	Val	Ala	Ala	Pro	Gln	Thr	Glu	Glu	Ala	Leu	Gly	
	65					70					75					
acc	gtg	cag	gct	gcc	ggc	gcg	ggc	gat	gag	cac	agc	gcc	gat	gta	gca	287
Thr	Val	Gln	Ala	Ala	Gly	Ala	Gly	Asp	Glu	His	Ser	Ala	Asp	Val	Ala	
80					85					90					95	
ctc	cag	cag	ctt	gac	cgg	gct	atc	gca	gag	cgt	cgt	gcc	cgg	cgc	aaa	335
Leu	Gln	Gln	Leu	Asp	Arg	Ala	Ile	Ala	Glu	Arg	Arg	Ala	Arg	Arg	Lys	
				100					105					110		
cgg	gag	cag	ctg	tca	tac	cag	gct	gcc	gcc	att	gca	gca	tca,	att	ggc	383
Arg	Glu	Gln	Leu	Ser	Tyr	Gln	Ala	Ala	Ala	Ile	Ala	Ala	Ser	Ile	Gly	
			115					120					125			
gtg	tca	ggc	att	gcc	atc	ttc	gcc	acc	tac	ctg	aga	ttt	gcc	atg	cac	431
Val	Ser	Gly	Ile	Ala	Ile	Phe	Ala	Thr	Tyr	Leu	Arg	Phe	Ala	Met	His	
		130					135					140				
atg	acc	gtg	ggc	ggc	gca	gtg	cca	tgg	ggt	gaa	gtg	gct	ggc	act	ctc	479
Met	Thr	Val	Gly	Gly	Ala	Val	Pro	Trp	Gly	Glu	Val	Ala	Gly	Thr	Leu	
	145					150					155					
ctc	ttg	gtg	gtt	ggt	ggc	gcg	ctc	ggc	atg	gag	atg	tat	gcc	cgc	tat	527
Leu	Leu	Val	Val	Gly	Gly	Ala	Leu	Gly	Met	Glu	Met	Tyr	Ala	Arg	Tyr	
160					165					170					175	
gca	cac	aaa	gcc	atc	tgg	cat	gag	tcg	cct	ctg	ggc	tgg	ctg	ctg	cac	575
Ala	His	Lys	Ala	Ile	Trp	His	Glu	Ser	Pro	Leu	Gly	Trp	Leu	Leu	His	
				180					185					190		

aag age cae cae aca eet ege act gga eee ttt gaa gee aae gae ttg	623
Lys Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu	
195 200 205	
ttt gca atc atc aat gga ctg ccc gcc atg ctc ctg tgt acc ttt ggc	671
Phe Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly	
210 215 220	
ttc tgg ctg ccc aac gtc ctg ggg gcg gcc tgc ttt gga gcg ggg ctg	719
Phe Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu	
225 230 235	
	•
ggc atc acg cta tac ggc atg gca tat atg ttt gta cac gat ggc ctg	767
Gly Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu	
240 245 250 255	
gtg cac agg cgc ttt ccc acc ggg ccc atc gct ggc ctg ccc tac atg	815
Val His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met	020
260 265 270	
aag ege etg aca gtg gee eac eag eta eac eac age gge aag tae ggt	863
Lys Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly	003
275 280 . 285	
203	
ggc gcg ccc tgg ggt atg ttc ttg ggt cca cag gag ctg cag cac att	911
Gly Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile	211
290 295 300	
253	
cca ggt gcg gcg gag gag gtg gag cga ctg gtc ctg gaa ctg gac tgg	050
Pro Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp	959
205	
305 310 315	
toc aag ogg tag ggtgoggaac caggoacgot ggtttcacac otcatgootg	1011
Ser Lys Arg	1011
320	
trataarata tarataraan sataratata assassabab abas assassab	
tgataaggtg tggctagagc gatgcgtgtg agacgggtat gtcacggtcg actggtctga	1071
tagggaatag categaggat statestant consent the same	
tggccaatgg catcggccat gtctggtcat cacgggctgg ttgcctgggt gaaggtgatg	1131
cacatcatca totocootto garagetos cacatatas	4404
cacatcatca tgtgcggttg gaggggctgg cacagtgtgg gctgaactgg agcagttgtc	1191

catattetat ttgtgggage tgagatgatg geatgettgg gatgtgeatg gateatggta 1311
gtgcagcaaa etatatteae etagggetgt tggtaggate aggtgaggee ttgeacattg 1371
catgatgtae tegteatggt gtgttggtga gaggatggat gtggatggat gtgtattete 1431
agaegtagae ettgaetgga ggettgateg agaggatggg eegtattett tgagagggga 1491
ggetegtgee agaaatggtg agtggatgae tgtgaegetg tacattgeag geaggtgaga 1551
tgeactgtet egattgtaaa atacatteag atgeaaaaaa aaaaaaaa aaaaaaa 1608

<210> 32

<211> 322

<212> PRT

<213> Haematococcus pluvialis

<400> 32

Thr Phe His Lys Pro Val Ser Gly Ala Ser Ala Leu Pro His Ile Gly

1 10 15

Pro Pro Pro His Leu His Arg Ser Phe Ala Ala Thr Thr Met Leu Ser 20 25 30

Lys Leu Gln Ser Ile Ser Val Lys Ala Arg Arg Val Glu Leu Ala Arg 35 40 45

Asp Ile Thr Arg Pro Lys Val Cys Leu His Ala Gln Arg Cys Ser Leu 50 55 60

Val Arg Leu Arg Val Ala Ala Pro Gln Thr Glu Glu Ala Leu Gly Thr 65 70 75 80

76/357

Val Gln Ala Ala Gly Ala Gly Asp Glu His Ser Ala Asp Val Ala Leu 85 90 95

Gln Gln Leu Asp Arg Ala Ile Ala Glu Arg Arg Ala Arg Arg Lys Arg 100 105 , 110

Glu Gln Leu Ser Tyr Gln Ala Ala Ile Ala Ala Ser Ile Gly Val 115 120 125

Ser Gly Ile Ala Ile Phe Ala Thr Tyr Leu Arg Phe Ala Met His Met 130 135 140

Thr Val Gly Gly Ala Val Pro Trp Gly Glu Val Ala Gly Thr Leu Leu 145 150 155 160

Leu Val Val Gly Gly Ala Leu Gly Met Glu Met Tyr Ala Arg Tyr Ala 165 170 175

His Lys Ala Ile Trp His Glu Ser Pro Leu Gly Trp Leu Leu His Lys 180 185 190

Ser His His Thr Pro Arg Thr Gly Pro Phe Glu Ala Asn Asp Leu Phe 195 200 205

Ala Ile Ile Asn Gly Leu Pro Ala Met Leu Leu Cys Thr Phe Gly Phe 210 215 220

Trp Leu Pro Asn Val Leu Gly Ala Ala Cys Phe Gly Ala Gly Leu Gly 225 230 235 240

Ile Thr Leu Tyr Gly Met Ala Tyr Met Phe Val His Asp Gly Leu Val

WO 2004/063359 PCT/EP2004/000099

250

255

77/357

245

His Arg Arg Phe Pro Thr Gly Pro Ile Ala Gly Leu Pro Tyr Met Lys
260 265 270

Arg Leu Thr Val Ala His Gln Leu His His Ser Gly Lys Tyr Gly Gly 275 280 285

Ala Pro Trp Gly Met Phe Leu Gly Pro Gln Glu Leu Gln His Ile Pro 290 295 300

Gly Ala Ala Glu Glu Val Glu Arg Leu Val Leu Glu Leu Asp Trp Ser 305 310 315 320

Lys Arg

<210> 33

<211> 528

<212> DNA

<213> Erwinia uredovora

<220>

<221> CDS

<222> (1)..(528)

<400> 33

atg ttg tgg att tgg aat gcc ctg atc gtt ttc gtt acc gtg att ggc 48

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly

1 5 10 15

atg gaa gtg att gct gca ctg gca cac aaa tac atc atg cac ggc tgg 96 Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp

20 25 30

ggt	tgg	gga	tgg	cat	ctt	tċa	cat	cat	gaa	ccg	cgt	aaa	ggt	gcg	ttt	144
Gly	Trp	Gly	Trp	His	Leu	Ser	His	His	Glu	Pro	Arg	Lys	Gly	Ala	Phe	
		35					40					45				
gaa	gtt	aac	gat	ctt	tat	gcc	gtg	gtt	ttt	gct	gca	tta	tcg	atc	ctg	192
Glu	Val	Asn	Asp	Leu	Tyr	Ala	Val	Val	Phe	Ala	Ala	Leu	Ser	Ile	Leu	
	50					55					60					
ctg	att	tat	ctg	ggc	agt	aca	gga	atg	tgg	ccg	ctc	cag	tgg	att	ggc	240
Leu	Ile	Tyr	Leu	Gly	Ser	Thr	Gly	Met	Trp	Pro	Leu	Gln	Trp	Ile	Gly	
65					70					75					80	
gca	ggt	atg	acg	gcg	tat	gga	tta	ctc	tat	ttt	atg	gtg	cac	gac	ggg	288
Ala	Gly	Met	Thr	Ala	Tyr	Gly	Leu	Leu	Tyr	Phe	Met	Val	His	Asp	Gly	
•				85					90					95		
ctg	gtg	cat	caa	cgt	tgg	cca	ttc	cgc	tat	att	cca	cgc	aag	ggc	tac	336
Leu	Val	His	Gln	Arg	Trp	Pro	Phe	Arg	Tyr	Ile	Pro	Arg	Lys	Gly	Tyr	
			100					105					110			
ctc	aaa	cgg	ttg	tat	atg	gcg	cac	cgt	atg	cat	cac	gcc	gtc	agg	ggc	384
Leu	Lys	Arg	Leu	Tyr	Met	Ala	His	Arg	Met	His	His	Ala	Val	Arg	Gly	
		115					120					125				
aaa	gaa	ggt	tgt	gtt	tct	ttt	ggc	ttc	ctc	tat	gcg	ccg	ccc	ctg	tca	432
Lys	Glu	Gly	Суз	Val	Ser	Phe	Gly	Phe	Leu	Tyr	Ala	Pro	Pro	Leu	Ser	
	130		•			135					140					
aaa	ctt	cag	gcg	acg	ctc	cgg	gaa	aga	cat	ggc	gct	aga	gcg	ggc	gct	480
Lys	Leu	Gln	Ala	Thr	Leu	Arg	Glu	Arg	His	Gly	Ala	Arg	Ala	Gly	Ala	
145					150					155					160	
gcc	aga	gat	gcg	cag	ggc	ggg	gag	gat	gag _.	ccc	gca	tcc	ggg	aag	taa	528
					Gly											
				165					170					175		

<210> 34

<211> 175

<212> PRT

<213> Erwinia uredovora

<400> 34

Met Leu Trp Ile Trp Asn Ala Leu Ile Val Phe Val Thr Val Ile Gly

1 5 10 15

Met Glu Val Ile Ala Ala Leu Ala His Lys Tyr Ile Met His Gly Trp
20 25 30

Gly Trp Gly Trp His Leu Ser His His Glu Pro Arg Lys Gly Ala Phe 35 40 45

Glu Val Asn Asp Leu Tyr Ala Val Val Phe Ala Ala Leu Ser Ile Leu 50 55 60

Leu Ile Tyr Leu Gly Ser Thr Gly Met Trp Pro Leu Gln Trp Ile Gly 65 70 75 80

Ala Gly Met Thr Ala Tyr Gly Leu Leu Tyr Phe Met Val His Asp Gly 85 90 95

Leu Val His Gln Arg Trp Pro Phe Arg Tyr Ile Pro Arg Lys Gly Tyr $100 \hspace{1.5cm} 105 \hspace{1.5cm} 110$

Leu Lys Arg Leu Tyr Met Ala His Arg Met His His Ala Val Arg Gly
115 120 125

Lys Glu Gly Cys Val Ser Phe Gly Phe Leu Tyr Ala Pro Pro Leu Ser 130 135 140

Lys Leu Gln Ala Thr Leu Arg Glu Arg His Gly Ala Arg Ala Gly Ala 145 150 155 160

WO 2004/063359 PCT/EP2004/000099 80/357

Ala Arg Asp Ala Gln Gly Gly Glu Asp Glu Pro Ala Ser Gly Lys

165 170 . 175

<210> 35

<211> 1520

<212> DNA

<213> Artificial

<220>

<223> Promotor

<400> 35

ctcgagtacc gaggcggaac ggcaggaatg tttccctctc ttttagaggg caattcttta 60 tecaatgtea tgttgatget agatatttet gtetettata ataaggegaa tacceatttt 120 tgaattgaag ttgagataaa aaaaaagggg gcccaatttg tcaacgccaa agagtcaagc 180 tttttctttg gctttagccg aacaatctaa gacttattgt ttttgaagat atttgacctt 240 ttctagatat tccttcaagt aaagcttttt tcgagttttt ttttttttc tttgtgaagg 300 atttattgtt attggtatcc atttttatt ggaagacaag ataagttaat attgattttg 360 cttaaagatt aaaaggaaat cagaaaacga caataaaaaa tgtaacggac aaactatggt 420 gtcgattata agtctaaatc cttaaaaaat gacaacgagt tgctttcctc tgaaaacaat 480 tettttgtet ttgcaagaaa ggtttetttt ttgtttgett geattaetta aacatcaaat 540 caaatgaaag gaataaagca gatttgaggg cgaataagga ttttctggtc aacaagatgt 600 gagtgacacc taaggaacta aatgccattc atttgtttta aaacgacatc aaagattgat 660 gatcaacagg attgagagag agaaaaagaa ctcgtgtcat ttatttctgt tgactgaaat 720 tttatattta gaaaaaatgt caaatctata gctttagcta tattacataa catttgaaat 780 aataataata aaaaaagaca cattagagac acttttcaaa ctctaaataa ctgtctataa 840

81/357

acacaaagaa aacaaagacc tctataacaa cttattagat ttttctcgta cttttgtcta 900 aagatgatgt attettgtta teccaeactt ettteatttg ttettgatge taetaaatat 960 acaaaatttc ttttttgcaa gagatattat tccaaaaatt ttcaaaaaga aatttttttc 1020 acaatagcag ttgatcgtgt aacccaaaga ggttctttgt tattttgcac ttccgctttg 1080 cggtgatgca tattcaaagt aatatatgga ataaacaacg tgtttaagca tgaaagaaag 1140 gaaacaaagg ccgctttgaa caaatgcata atatttcaga caaaaatgat ctaaagcaag 1200 cagtaaatca aacaagaaac attgctgatt cgcgttagaa aacgataaaa gtctaataag 1260 ccactaagta tacttcaatg aactttttgt atgcttatgg tccaatcaga ccaataattt 1320 gtgaccattc ctgaggtggc tttggtgatg cggaaacaga aaaaaatttt ctcaccaatc 1380 gatttaaaaa acaatttctg ctttgaacca aaactttttt tttctcttta atcattaact 1440 ttatcaagta tgtacctacc ctcaaagtcc tcactcaagc acaattatgc taacattgtt 1500 1520 ccaccttctc tttagaaatg

<210> 36

<211> 16245

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

WO 2004/063359 . PCT/EP2004/000099

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 36

cegggetggt tgecetegee getgggetgg eggeegteta tggeeetgea aacgegeeag 60 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 120 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180 ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240 cagectegea aateggegaa aaegeetgat tttaegegag ttteecacag atgatgtgga 300 caagectggg gataagtgee etgeggtatt gacaettgag gggegegaet aetgacagat 360 gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 420 tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgccccc 600 cttetegaac ceteceggee egetaacgeg ggeeteecat eececeaggg getgegeece 660 tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720 atcggggcag taacgggatg ggcgatcagc ccgagcgcga cgcccggaag cattgacgtg 780 ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840 ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900 gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960 ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020

WO 2004/063359 PCT/EP2004/000099 83/357

acgagaattg gaccttt	aca gaattactct	atgaagegee	atatttaaaa	agctaccaag	1080
acgaagagga tgaagag	gat gaggaggcag	g attgccttga	atatattgac	aatactgata	1140
agataatata totttta	tat agaagatato	: gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg cgcttat	caa tatatctata	gaatgggcaa	agcataaaaa	cttgcatgga	1260
ctaatgcttg aaaccca	gga caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa cttattg	ata gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg.attttga	gaa cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt tatgccg	ctc aattegetge	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc gggattc	ata cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca ggctcat	aag acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg tcttccg	gag actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat agcccca	ctg ttcgtccatt	tccgcgcaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg aggttac	cga ctgcggcctg	agtttttaa	gtgacgtaaa	atcgtgttga	1800
ggccaacgcc cataatg	egg getgttgeee	ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg gtgcgtad	ccg ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg agagcaga	aga tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt agctgaad	cag gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca tcatacac	ta aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc cgtcgata	act atgttatacg	ccaactttga	aaacaacttt	gaaaaagctg	2160
ttttctggta tttaaggt	tt tagaatgcaa	ggaacagtga	attggagttc	gtcttgttat	2220
aattagcttc ttggggta	itc tttaaatact	gtagaaaaga	ggaaggaaat	aataaatggc	2280

WO 2004/063359 PCT/EP2004/000099 84/357

			3.7007			
taaaatgaga	atatcaccgg	aattgaaaaa	actgatcgaa	aaataccgct	gcgtaaaaga	2340
tacggaagga	atgtctcctg	ctaaggtata	taagctggtg	ggagaaaatg	aaaacctata	2400
tttaaaaatg	acggacagcc	ggtataaagg	gaccacctat	gatgtggaac	gggaaaagga	2460
catgatgcta	tggctggaag	gaaagctgcc	tgttccaaag	gtcctgcact	ttgaacggca	2520
tgatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac	ttactgaata	acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt	aaagatccgc	gcgagctgta	tgatttttta	aagacggaaa	agcccgaaga	2820
ggaacttgtc	ttttcccacg	gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc	tttattgatc	ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc	cggtcgatca	gggaggatat	cggggaagaa	cagtatgtcg	agctattttt	3000
tgacttactg	gggatcaagc	ctgattggga	gaaaataaaa	tattatattt	tactggatga	3060
attgttttag	tacctagatg	tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tetteegeat	caagtgtttt	ggeteteagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt	attcgtgcag	ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg	gaccgacttc	attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa	tcaggaataa	gggcacattg	ccccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc	cgaggatgcc	gaaaccatcg	caagccgcac	cgtcatgcgt	gcgccccgcg	3480
aaaccttcca	gtccgtcggc	tcgatggtcc	agcaagctac	ggccaagatc	gagegegaea	3540

WO 2004/063359 PCT/EP2004/000099 85/357

gcgtgcaact ggctccccct gccctgcccg cgccatcggc cgccgtggag cgttcgcgtc 3600 gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660 tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720 agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780 3840 tegatattgc geegtggceg gacacgatgc gagegatgcc aaacgacacg geeegetetg ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900 tecaegteaa caaggaegtg aagateaect acaeeggegt egagetgegg geegaegatg 3960 acgaactggt gtggcagcag gtgttggagt acgcgaagcg cacccctatc ggcgagccga 4020 4080 tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4140 acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg 4200 accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4260 gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320 4380 ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440 cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4500 tgcattgcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4560 ctttactggc atttcaggaa caagegggca ctgctcgacg cacttgcttc gctcagtatc 4620 4680 gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4740 tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat 4800 ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga

			00,007			
gaaaaagccc	atggaggcgt	tcgctgaacg	gttgcgagat	gccgtggcat	teggegeeta	4860
catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagegaggee	gaggggtcgc	4980
cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	tccgacagat	5040
tccaacggga	atctggtgga	tgcgcatctt	catcctcggc	gcacttaata	tttcgctatt	5100
ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160
cgctgtgcag	ccgctgatgg	tcgtgttcat	ctctgccgct	ctgctaggta	gcccgatacg	5220
attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	tggtgttgac	5280
accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	cggtttccat	5340
ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	tcacctttac	5400
cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	tgtttgatcc	5460
gccaatcccg	atgcctacag	gaaccaatgt	tctcggcctg	gcgtggctcg	gcctgatcgg	5520
agcgggttta	acctacttcc	tttggttccg	ggggatctcg	cgactcgaac	ctacagttgt	5580
ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	tgcatcaggc	5640
cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	ataggggagt	5700
tgatatcgtc	aacgttcact	tctaaagaaa	tagcgccact	cagcttcctc	ageggettta	5760
tccagcgatt	tcctattatg	teggcatagt	tctcaagatc	gacagcctgt	cacggttaag	5820
cgagaaatga	ataagaaggc	tgataattcg	gatetetgeg	agggagatga	tatttgatca	5880
caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	gatcatccgt	5940
gtttcaaacc	cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	atgagcaaag	6000
tctgccgcct	tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	ctgcctgtat	6060

WO 2004/063359 PCT/EP2004/000099 87/357

cgagtggtga ttttgtgccg	agctgccggt	cggggagctg	ttggctggct	ggtggcagga	6120
tatattgtgg tgtaaacaaa	ttgacgctta	gacaacttaa	taacacattg	cggacgtttt	6180
taatgtactg gggtggtttt	tcttttcacc	agtgagacgg	gcaacagctg	attgcccttc	6240
accgcctggc cctgagagag	ttgcagcaag	cggtccacgc	tggtttgccc	cagcaggcga	6300
aaatcctgtt tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	6360
ccgagatagg gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420
actccaacgt caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480
cacccaaatc aagttttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	6540
ggagcccccg atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600
agaaagcgaa aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720
tgggtaacgc cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780
ctcggtaccc ggggatcttt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840
cgaggagcct cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900
gctcattaag agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac gctcttctga	tccagtcgat	catccgctga	aggegettte	gaatctggtt	7020
aagatecaeg tettegggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080
aacagettte teagecaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140
aagaactgga ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320

WO 2004/063359 PCT/EP2004/000099 88/357

tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
tcgaaataca	tccggatgtc	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560
tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaacatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ggtggagagc	ttataccgag	ctcccaaatc	7920
tgtccagatc	atggttgacc	ggtgcctgga	tcttcctata	gaatcatcct	tattcgttga	7980
cctagctgat	tctggagtga	cccagagggt	catgacttga	gcctaaaatc	cgccgcctcc	8040
accatttgta	gaaaaatgtg	acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100
tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataagc	cactggccag	8160
acagctctgg	cggctctgag	gtgcagtgga	tgattattaa	tccgggaccg	gccgcccctc	8220
cgccccgaag	tggaaaggct	ggtgtgcccc	tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatageegt	cggcgaaata	gcatgccatt	aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	agattcacga	gatagtacct	tctccgaagt	aggtagagcg	agtacccggc	8460
gcgtaagctc	cctaattggc	ccatccggca	tctgtagggc	gtccaaatat	cgtgcctctc	8520
ctgctttgcc	cggtgtatga	aaccggaaag	gccgctcagg	agctggccag	cggcgcagac	8580

WO 2004/063359 PCT/EP2004/000099 89/357

cgggaacaca	agctggcagt	cgacccatcc	ggtgctctgc	actcgacctg	ctgaggtccc	8640
tcagtccctg	gtaggcagct	ttgccccgtc	tgtccgcccg	gtgtgtcggc	ggggttgaca	8700
aggtcgttgc	gtcagtccaa	catttgttgc	catattttcc	tgctctcccc	accagctgct	8760
cttttcttt	ctctttcttt	tcccatcttc	agtatattca	tcttcccatc	caagaacctt	8820
tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	gcccattcgg	9360
accgcaagga	atcggtcaat	acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	tegegeagge	9480
tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	tcgtgcacgc	9540
ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	cggagcttgc	9720
aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840

WO 2004/063359 PCT/EP2004/000099 90/357

ccgatccgga	gccgggactg	tcgggcgtac	acaaatcgcc	cgcagaagcg	cggccgtctg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960
gagggcaaag	gaatagagta	gatgccgacc	gcgggatcga	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080
ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttcgggttta	10200
cctcttccag	atacagetca	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagagc	tattttcatt	ttcgggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380
tccacgcgac	tatatatttg	tctctaattg	tactttgaca	tgctcctctt	ctttactctg	10440
atagcttgac	tatgaaaatt	ccgtcaccag	cncctgggtt	cgcaaagata	attgcatgtt	10500
tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
canttcctac	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactctc	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tetecaeteg	acctgcaggc	atgcaagctt	aatctataca	10800
atgctccata	gactcacatt	gatattgtcg	aagatttcga	tgctgactta	gtagagcaac	10860
tacaaaagtt	agcagagaag	catgatttct	taatctttga	agaccgcaag	tttgcagata	10920
tcggtatgtg	aattctatct	atttttttc	tgatgtgtgc	atggatgact	catgatcata	10980
ttcttaggta	atactgtcaa	gcatcaatat	ggcaagggcg	tttacaagat	tgcttcttgg	11040
tctcatatta	ctaatgctca	cacagttcct	ggagaaggta	ttatcaaggg	acttgccgaa	11100

WO 2004/063359 PCT/EP2004/000099 91/357

gteggeetee etettggteg tggettgett ttgctageag aaatgteate teaaggtgea 11160 ttaactaagg gtatttacac tgccgaatct gtcaatatgg ctcgccgcaa caaagatttc 11220 gtttttggct ttattgcaca acacaaaatg aatcagtatg atgatgagga ttttgttgtc 11280 atgtcgcctg aagcttggcg taatcatggt catagctgtt tcctgtgtga aattgttatc 11340 cgctcacaat tccacacaac atacgagccg gaagcataaa gtgtaaagcc tggggtgcct 11400 aatgagtgag ctaactcaca ttaattgcgt tgcgctcact gcccgctttc cagtcgggaa 11460 acctgtcgtg ccagctgcat taatgaatcg gccaacgcgc ggggagaggc ggtttgcgta 11520 ttgggccaaa gacaaaaggg cgacattcaa ccgattgagg gagggaaggt aaatattgac 11580 ggaaattatt cattaaaggt gaattatcac cgtcaccgac ttgagccatt tgggaattag 11640 agccagcaaa atcaccagta gcaccattac cattagcaag gccggaaacg tcaccaatga 11700 aaccatcgat agcagcaccg taatcagtag cgacagaatc aagtttgcct ttagcgtcag 11760 actgtagcgc gttttcatcg gcattttcgg tcatagcccc cttattagcg tttgccatct 11820 tttcataatc aaaatcaccg gaaccagagc caccaccgga accgcctccc tcagagccgc 11880 cacecteaga acegecacec teagageeac eacecteaga geegecacea gaaceaceac 11940 cagageegee gecageattg acaggaggee egatetagta acatagatga cacegegege 12000 gataatttat cctagtttgc gcgctatatt ttgttttcta tcgcgtatta aatgtataat 12060 tgcgggactc taatcataaa aacccatctc ataaataacg tcatgcatta catgttaatt 12120 attacatgct taacgtaatt caacagaaat tatatgataa tcatcgcaag accggcaaca 12180 ggattcaatc ttaatgaaact ttattgccaa atgtttgaac gatcggggat catccgggtc 12240 tgtggcggga actccacgaa aatatccgaa cgcagcaaga tatcgcggtg catctcggtc 12300 ttgcctgggc agtcgccgcc gacgccgttg atgtggacgc cgggcccgat catattgtcg 12360

WO 2004/063359 PCT/EP2004/000099 92/357

			12/03/			
ctcaggatcg	g tggcgttgtg	cttgtcggcc	gttgctgtcg	taatgatato	ggcaccttcg	12420
accgcctgtt	ccgcagagat	cccgtgggcg	aagaactcca	gcatgagato	cccgcgctgg	12480
aggateatec	agccggcgtc	ccggaaaacg	attccgaagc	ccaacctttc	atagaaggcg	12540
gcggtggaat	: cgaaatctcg	tgatggcagg	ttgggcgtcg	cttggtcggt	catttcgaac	12600
cccagagtcc	cgctcagaag	aactcgtcaa	gaaggcgata	gaaggcgatg	cgctgcgaat	12660
cgggagcggc	gataccgtaa	agcacgagga	agcggtcagc	ccattcgccg	ccaagctctt	12720
cagcaatatc	acgggtagcc	aacgctatgt	cctgatagcg	gtccgccaca	cccagccggc	12780
cacagtcgat	gaatccagaa	aagcggccat	tttccaccat	gatattcggc	aagcaggcat	12840
cgccatgggt	cacgacgaga	tcatcgccgt	cgggcatgcg	egcettgage	ctggcgaaca	12900
gttcggctgg	cgcgagcccc	tgatgctctt	cgtccagatc	atcctgatcg	acaagaccgg	12960
cttccatccg	agtacgtgct	cgctcgatgc	gatgtttcgc	ttggtggtcg	aatgggcagg	13020
tagccggatc	aagcgtatgc	agccgccgca	ttgcatcagc	catgatggat	actttctcgg	13080
caggagcaag	gtgagatgac	aggagateet	gccccggcac	ttcgcccaat	agcagccagt	13140
cccttcccgc	ttcagtgaca	acgtcgagca	cagctgcgca	aggaacgccc	gtcgtggcca	13200
gccacgatag	ccgcgctgcc	tcgtcctgca	gttcattcag	ggcaccggac	aggtcggtct	13260
tgacaaaaag	aaccgggcgc	ccctgcgctg	acagccggaa	cacggcggca	tcagagcagc	13320
cgattgtctg	ttgtgcccag	tcatagccga	atageetete	cacccaagcg	gccggagaac	13380
ctgcgtgcaa	tccatcttgt	tcaatcatgc	gaaacgatcc	agatccggtg	cagattattt	13440
ggattgagag	tgaatatgag	actctaattg	gataccgagg	ggaatttatg	gaacgtcagt	13500
ggagcatttt	tgacaagaaa	tatttgctag	ctgatagtga	ccttaggcga	cttttgaacg	13560
cgcaataatg	gtttctgacg	tatgtgctta	gctcattaaa	ctccagaaac	ccgcggctga	13620

WO 2004/063359 PCT/EP2004/000099 93/357

gtggctcctt	caacgttgcg	gttctgtcag	ttccaaacgt	aaaacggctt	gtcccgcgtc	13680
atcggcgggg	gtcataacgt	gactccctta	attctccgct	catgatcaga	ttgtcgtttc	13740
ccgccttcag	tttaaactat	cagtgtttga	caggatatat	tggcgggtaa	acctaagaga	13800
aaagagcgtt	tattagaata	atcggatatt	taaaagggcg	tgaaaaggtt	tatccgttcg	13860
tccatttgta	tgtgcatgcc	aaccacaggg	ttccccagat	ctggcgccgg	ccagcgagac	13920
gagcaagatt	ggccgccgcc	cgaaacgatc	cgacagcgcg	cccagcacag	gtgcgcaggc	13980
aaattgcacc	aacgcataca	gegecageag	aatgccatag	tgggcggtga	cgtcgttcga	14040
gtgaaccaga	tegegeagga	ggcccggcag	caccggcata	atcaggccga	tgccgacagc	14100
gtcgagcgcg	acagtgctca	gaattacgat	caggggtatg	ttgggtttca	cgtctggcct	14160
ccggaccagc	ctccgctggt	ccgattgaac	gcgcggattc	tttatcactg	ataagttggt	14220
ggacatatta	tgtttatcag	tgataaagtg	tcaagcatga	caaagttgca	gccgaataca	14280
gtgatccgtg	ccgccctgga	cctgttgaac	gaggtcggcg	tagacggtct	gacgacacgc	14340
aaactggcgg	aacggttggg	ggttcagcag	ceggegettt	actggcactt	caggaacaag	14400
cgggcgctgc	tcgacgcact	ggccgaagcc	atgctggcgg	agaatcatac	gcattcggtg	14460
ccgagagccg	acgacgactg	gcgctcattt	ctgatcggga	atgcccgcag	cttcaggcag	14520
gegetgeteg	cctaccgcga	tggcgcgcgc	atccatgccg	gcacgcgacc	gggcgcaccg	14580
cagatggaaa	cggccgacgc	gcagcttcgc	ttcctctgcg	aggcgggttt	ttcggccggg	14640
gacgccgtca	atgcgctgat	gacaatcagc	tacttcactg	ttggggccgt	gcttgaggag	14700
caggccggcg	acagegatge	cggcgagcgc	ggcggcaccg	ttgaacaggc	tccgctctcg	14760
ccgctgttgc	gggccgcgat	agacgccttc	gacgaagccg	gtccggacgc	agcgttcgag	14820
cagggactcg	cggtgattgt	cgatggattg	gcgaaaagga	ggctcgttgt	caggaacgtt	14880

WO 2004/063359 PCT/EP2004/000099 94/357

gaaggaccga gaaagggtga cgattgatca ggaccgctgc cggagcgcaa cccactcact 14940 acagcagage catgtagaca acateceete eccettteca eegegteaga egeeegtage agcccgctac gggctttttc atgccctgcc ctagcgtcca agcctcacgg ccgcgctcgg cetetetgge ggeettetgg egetetteeg etteeteget eactgacteg etgegetegg tegttegget geggegageg gtateagete aeteaaagge ggtaataegg ttatecaeag aatcagggga taacgcagga aagaacatgt gagcaaaagg ccagcaaaag gccaggaacc gtaaaaaggc cgcgttgctg gcgtttttcc ataggctccg ccccctgac gagcatcaca 15300 aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga taccaggcgt ttccccctgg aagetecete gtgegetete etgtteegae eetgeegett aeeggataee tgtccgcctt tctcccttcg ggaagcgtgg cgcttttccg ctgcataacc ctgcttcggg 15480 gtcattatag cgatttttc ggtatatcca tcctttttcg cacgatatac aggattttgc 15540 caaagggttc gtgtagactt tccttggtgt atccaacggc gtcagccggg caggataggt 15600 gaagtaggcc caccegegag egggtgttcc ttetteactg teeettatte geacetggeg 15660 gtgctcaacg ggaatcctgc tctgcgaggc tggccggcta ccgccggcgt aacagatgag 15720 ggcaagcgga tggctgatga aaccaagcca accaggaagg gcagcccacc tatcaaggtg 15780 tactgccttc cagacgaacg aagagcgatt gaggaaaagg cggcggcggc cggcatgagc 15840 ctgtcggcct acctgctggc cgtcggccag ggctacaaaa tcacgggcgt cgtggactat 15900 gagcacgtcc gcgagctggc ccgcatcaat ggcgacctgg gccgcctggg cggcctgctg 15960 aaactctggc tcaccgacga cccgcgcacg gcgcggttcg gtgatgccac gatcctcgcc 16020 ctgctggcga agatcgaaga gaagcaggac gagcttggca aggtcatgat gggcgtggtc 16080 cgcccgaggg cagagccatg acttttttag ccgctaaaac ggccgggggg tgcgcgtgat 16140 tgccaagcac gtccccatgc gctccatcaa gaagagcgac ttcgcggagc tggtgaagta 16200

catcaccgac gagcaaggca agaccgagcg cctttgcgac gctca 16245

<210> 37

<211> 17877

<212> DNA

<213> Artificial

<220>

<223> Promotor

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 37

cegggetggt tgeeetegee getgggetgg eggeegteta tggeeetgea aaegegeeag 60

aaaegeegte gaageegtgt gegagacaee geggeegeeg gegttgtgga taeetegegg 120

aaaaettgge ceteaetgae agatgagggg eggaegttga eaettgaggg geegaeteae 180

ceggegegge gttgacagat gaggggeagg etegattteg geeggegaeg tggagetgge 240

cageetegea aateggegaa aaegeetgat tttaegegag ttteeeacag atgatgtgga 300

caageetggg gataagtgee etgeggtatt gacaettgag gggegegaet aetgacagat

360

WO 2004/063359 PCT/EP2004/000099 96/357

			70/351			
gaggggcgcg	atccttgaca	cttgaggggc	agagtgctga	cagatgaggg	gcgcacctat	420
tgacatttga	ggggctgtcc	acaggcagaa	aatccagcat	ttgcaagggt	ttccgcccgt	480
ttttcggcca	ccgctaacct	gtcttttaac	ctgcttttaa	accaatattt	ataaaccttg	540
tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcģcac	gccgaagggg	ggtgccccçc	600
cttctcgaac	cctcccggcc	cgctaacgcg	ggcctcccat	cccccaggg	gctgcgcccc	660
teggeegega	acggcctcac	cccaaaaatg	gcagcgctgg	cagtecttge	cattgccggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	eggeggeetg	840
ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaattttta	ccttgggcat	tettggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg	gacctttaca	gaattactct	atgaagcgcc	atatttaaaa	agctaccaag	1080
acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	aatactgata	1140
agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg	cgcttatcaa	tatatctata	gaatgggcaa	agcataaaaa	cttgcatgga	1260
ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgccgctc	aattcgctgc	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620

WO 2004/063359 PCT/EP2004/000099 97/357

9	gcaacaaccg	tetteeggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
ç	gccccgacat	agccccactg	ttcgtccatt	tccgcgcaga	cgatgacgtc	actgcccggc	1740
t	tgtatgcgcg	aggttaccga	ctgcggcctg	agttttttaa	gtgacgtaaa	atcgtgttga	1800
ç	ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
t	gattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
t	cacggcagtg	agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
C	cccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
č	aaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
έ	aatcggctc	cgtcgatact	atgttatacg	ccaactttga	aaacaacttt	gaaaaagctg	2160
t	tttctggta	tttaaggttt	tagaatgcaa	ggaacagtga	attggagttc	gtcttgttat	2220
ε	aattagcttc	ttggggtatc	tttaaatact	gtagaaaaga	ggaaggaaat	aataaatggc	2280
t	taaaatgaga	atatcaccgg	aattgaaaaa	actgatcgaa	aaataccgct	gcgtaaaaga	2340
t	acggaagga	atgtctcctg	ctaaggtata	taagctggtg	ggagaaaatg	aaaacctata	2400
t	ttaaaaatg	acggacagcc	ggtataaagg	gaccacctat	gatgtggaac	gggaaaagga	2460
c	atgatgcta	tggctggaag	gaaagetgee	tgttccaaag	gtcctgcact	ttgaacggca	2520
t	gatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
t	gaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
t	cactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
ć	attggattac	ttactgaata	acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
c	cactccattt	aaagatccgc	gcgagctgta	tgatttttta	aagacggaaa	agcccgaaga	2820
ç	ggaacttgtc	ttttcccacg	gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880

WO 2004/063359 PCT/EP2004/000099 98/357

agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggt atgacattgc cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattttt 3000 tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggatga 3060 attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120 tetteegeat caagtgtttt ggeteteagg cegaggeeca eggeaagtat ttgggeaagg 31.80 ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga 3240 3300 cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgcaag 3360 gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420 3480 ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3540 aaaccttcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc gagcgcgaca gegtgeaact ggeteecet geeetgeeeg egeeategge egeegtggag egttegegte 3600 gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660 3720 tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3780 agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt tegatattge geogtggeeg gacacgatge gagegatgee aaacgacaeg geoegetetg 3840 ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt 3900 tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960 4020 acgaactggt gtggcagcag gtgttggagt acgcgaagcg cacccctatc ggcgagccga tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080 4140 acacgaaggc cgaggaatgc ctgtcgcgcc tacaggcgac ggcgatgggc ttcacgtccg

WO 2004/063359 PCT/EP2004/000099 99/357

accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg 4200 gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg 4260 gegaceacta caegaaatte atatgggaga agtacegeaa getgtegeeg aeggeeegae 4320 ggatgttcga ctatttcagc tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc 4380 gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag 4440 4500 cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg 4560 tgcattgcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg 4620 ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat 4680 4740 tgtgattaag geteagatte gaeggettgg ageggeegae gtgeaggatt teegegagat ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga 4800 gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta 4920 categacgge gagateattg ggetgteggt etteaaacag gaggacggee ceaaggacge 4980 tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc 5040 cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100 ctggagcttg ttgtttattt cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160 cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220 5280 attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tggtgttgac 5340 accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5400 ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac

cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc gecaateceg atgeetacag gaaceaatgt teteggeetg gegtggeteg geetgategg 5520 agegggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580 ttccttactg ggctttctca gccccagatc tggggtcgat cagccgggga tgcatcaggc 5640 cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700 tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760 tecagegatt tectattatg teggeatagt teteaagate gacageetgt caeggttaag 5820 cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880 caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940 gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000 tetgeegeet tacaaegget etecegetga egeegteeeg gaetgatggg etgeetgtat 6060 cgagtggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6120 tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180 taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240 accgcctggc cctgagagag ttgcagcaag cggtccacgc tggtttgccc cagcaggcga 6300 aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6360 ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg 6420 actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6540 ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6600

WO 2004/063359 PCT/EP2004/000099 101/357

agaaagcgaa	aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt	cgctattacg	ccagetggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720
tgggtaacgc	cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780
ctcggtaccc	ggggatcttt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840
cgaggagcct	cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900
gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac	gctcttctga	tccagtcgat	catccgctga	aggcgctttc	gaatctggtt	7020
aagatccacg	tcttcgggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140
aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag	aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320
tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
tcgaaataca	tccggatgtc	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560
tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaacatct	cctggataaa	7860

WO 2004/063359 PCT/EP2004/000099 102/357

ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatc 7920 tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980 cctagctgat tctggagtga cccagagggt catgacttga gcctaaaatc cgccgcctcc 804Ò accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc 8100 tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160 acagetetgg eggetetgag gtgeagtgga tgattattaa teegggaeeg geegeeete 8220 cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaatctatt gcatcatcgg 8280 agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt 8340 gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400 atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtacccggc 8460 gegtaagete eetaattgge ceateeggea tetgtaggge gtecaaatat egtgeetete 8520 ctgctttgcc cggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580 cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640 teagteeetg gtaggeaget ttgeeeegte tgteegeeeg gtgtgtegge ggggttgaea aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760 cttttcttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt 8820 tatttcccct aagtaagtac tttgctacat ccatactcca tccttcccat cccttattcc 8880 tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgc 8940 ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga 9000 aaagttegae agegteteeg acetgatgea geteteggag ggegaagaat etegtgettt 9060 cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt 9120

WO 2004/063359 PCT/EP2004/000099 103/357

ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat	gegategetg	cggccgatct	tagccagacg	agcgggttcg	gcccattcgg	9360
		acactacatg				9420
		ctgtgatgga				9480
		gggccgagga				9540
		tcctgacgga				9600
		attcccaata				9660
		agcagacgcg				9720
9199119911	cycacygage	ageagaegeg	ozacoogag	0334330400	0994900050	J, L
aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga	gccgggactg	tcgggcgtac	acaaatcgcc	cgcagaagcg	cggccgtctg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960
gagggcaaag	gaatagagta	gatgccgacc	gcgggatcga	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080
ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttcgggttta	10200
cctcttccag	atacagctca	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagagc	tattttcatt	ttcġggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380

WO 2004/063359 PCT/EP2004/000099 104/357

tccacgcgac	tatatatttg	tctctaattg	tactttgaca	tgctcctctt	ctttactctg	10440
atagcttgac	tatgaaaatt	ccgtcaccag	cncctgggtt	cgcaaagata	attgcatgtt	10500
tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
canttcctac	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactctc	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tctccactcg	acctgcaggc	atgcaagctt	ttttcgagtt	10800
ttttttttt	ttctttgtga	aggatttatt	gttattggta	tccattttt	attggaagac	10860
aagataagtt	aatattgatt	ttgcttaaag	attaaaagga	aatcagaaaa	cgacaataaa	10920
aaatgtaacg	gacaaactat	ggtgtcgatt	ataagtctaa	atccttaaaa	aatgacaacg	10980
agttgctttc	ctctgaaaac	aattcttttg	tetttgcaag	aaaggtttct	tttttgtttg	11040
cttgcattac	ttaaacatca	aatcaaatga	aaggaataaa	gcagatttga	gggcgaataa	11100
ggattttctg	gtcaacaaga	tgtgagtgac	acctaaggaa	ctaaatgcca	ttcatttgtt	11160
ttaaaacgac	atcaaagatt	gatgatcaac	aggattgaga	gagagaaaaa	gaactcgtgt	11220
catttatttc	tgttgactga	aattttatat	ttagaaaaaa	tgtcaaatct	atagctttag	11280
ctatattaca	taacatttga	aataataata	ataaaaaaag	acacattaga	gacacttttc	11340
aaactctaaa	taactgtcta	taaacacaaa	gaaaacaaag	acctctataa	caacttatta	11400
gatttttctc	gtacttttgt	ctaaagatga	tgtattcttg	ttatcccaca	cttctttcat	11460
ttgttcttga	tgctactaaa	tatacaaaat	ttctttttg	caagagatat	tattccaaaa	11520
attttcaaaa	agaaattttt	ttcacaatag	cagttgatcg	tgtaacccaa	agaggttett	11580
tgttattttg	cacttccgct	ttgcggtgat	gcatattcaa	agtaatatat	ggaataaaca	11640

WO 2004/063359 PCT/EP2004/000099 105/357

acgtgtttaa gcatgaaaga aaggaaacaa aggccgcttt gaacaaatgc ataatatttc 11700 agacaaaaat gatctaaagc aagcagtaaa tcaaacaaga aacattgctg attcgcgtta 11760 gaaaacgata aaagtctaat aagccactaa gtatacttca atgaactttt tgtatgctta 11820 tggtccaatc agaccaataa tttgtgacca ttcctgaggt ggctttggtg atgcggaaac 11880 agaaaaaaat tttctcacca atcgatttaa aaaacaattt ctgctttgaa ccaaaacttt 11940 ttttttctct ttaatcatta actttatcaa gtatgtacct accctcaaag tcctcactca 12000 agcacaatta tgctaacatt gttccacctt ctctttagaa atgctgtcga agctgcagtc 12060 aatcagegte aaggeeegee gegttgaaet ageeegegae atcaegegge eeaaagtetg 12120 cetgeatget cageggtget egttagtteg getgegagtg geageaceae agaeagagga 12180 ggcgctggga accgtgcagg ctgccggcgc gggcgatgag cacagcgccg atgtagcact 12240 ccagcagett gacegggeta tegeagageg tegtgeeegg egeaaaeggg ageagetgte 12300 ataccagget geogecattg cagcatcaat tggegtgtca ggeattgeca tettegecae 12360 ctacctgaga tttgccatgc acatgaccgt gggcggcgca gtgccatggg gtgaagtggc 12420 tggcactete etettggtgg ttggtggege geteggeatg gagatgtatg eeegetatge 12480 acacaaagee atetggeatg agtegeetet gggetggetg etgeacaaga geeaceacae 12540 acctegeact ggaccetttg aagecaacga ettgtttgca atcateaatg gactgeeege 12600 catgeteetg tgtacetttg gettetgget geccaaegte etgggggegg cetgetttgg 12660 agcggggctg ggcatcacgc tatacggcat ggcatatatg tttgtacacg atggcctggt 12720 geacaggege ttteceaceg ggeecatege tggeetgeee tacatgaage geetgacagt 12780 ggcccaccag ctacaccaca gcggcaagta cggtggcgcg ccctggggta tgttcttggg 12840 tccacaggag ctgcagcaca ttccaggtgc ggcggaggag gtggagcgac tggtcctgga 12900

WO 2004/063359 PCT/EP2004/000099 106/357

actggactgg tccaagcggt agaagcttgg cgtaatcatg gtcatagctg tttcctgtgt 12960 gaaattgtta teegeteaca atteeacaca acatacgage eggaageata aagtgtaaag 13020 cctggggtgc ctaatgagtg agctaactca cattaattgc gttgcgctca ctgcccgctt 13080 tccagtcggg aaacctgtcg tgccagctgc attaatgaat cggccaacgc gcggggagag 13140 gcggtttgcg tattgggcca aagacaaaag ggcgacattc aaccgattga gggagggaag 13200 gtaaatattg acggaaatta ttcattaaag gtgaattatc accgtcaccg acttgagcca 13260 tttgggaatt agagccagca aaatcaccag tagcaccatt accattagca aggccggaaa 13320 cgtcaccaat gaaaccatcg atagcagcac cgtaatcagt agcgacagaa tcaagtttgc 13380 ctttagcgtc agactgtagc gcgttttcat cggcattttc ggtcatagcc cccttattag 13440 cgtttgccat cttttcataa tcaaaatcac cggaaccaga gccaccaccg gaaccgcctc 13500 ceteagagee gecaceetea gaacegecae ecteagagee accaceetea gageegecae 13560 cagaaccacc accagageeg cegecageat tgacaggagg ceegatetag taacatagat 13620 gacaccgcgc gcgataattt atcctagttt gcgcgctata ttttgttttc tatcgcgtat 13680 taaatgtata attgcgggac tctaatcata aaaacccatc tcataaataa cgtcatgcat 13740 tacatgttaa ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca 13800 agaccggcaa caggattcaa tcttaagaaa ctttattgcc aaatgtttga acgatcgggg 13860 atcatccggg tctgtggcgg gaactccacg aaaatatccg aacgcagcaa gatatcgcgg 13920 tgcatctcgg tcttgcctgg gcagtcgccg ccgacgccgt tgatgtggac gccgggcccg 13980 atcatattgt cgctcaggat cgtggcgttg tgcttgtcgg ccgttgctgt cgtaatgata 14040 teggeacett egacegeetg tteegeagag atecegtggg egaagaacte cageatgaga 14100

WO 2004/063359 PCT/EP2004/000099 107/357

teccegeget ggaggateat ecageeggeg teceggaaaa egatteegaa geecaacett 14160 tcatagaagg cggcggtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg 14220 gtcatttcga accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga 14280 tgcgctgcga atcgggagcg gcgataccgt aaagcacgag gaagcggtca gcccattcgc 14340 cgccaagctc ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggtccgcca 14400 cacccagccg gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg 14460 gcaagcaggc atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgccttga 14520 gcctggcgaa cagttcggct ggcgcgagcc cctgatgctc ttcgtccaga tcatcctgat 14580 cgacaagacc ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggt 14640 cgaatgggca ggtagccgga tcaagcgtat gcagccgccg cattgcatca gccatgatgg 14700 atactttctc ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgccca 14760 atagcageca gtecetteee getteagtga caacgtegag cacagetgeg caaggaacge 14820 ccgtcgtggc cagccacgat agccgcgctg cctcgtcctg cagttcattc agggcaccgg 14880 acaggtcggt cttgacaaaa agaaccgggc gcccctgcgc tgacagccgg aacacggcgg 14940 catcagagca gccgattgtc tgttgtgccc agtcatagcc gaatagcctc tccacccaag 15000 cggccggaga acctgcgtgc aatccatctt gttcaatcat gcgaaacgat ccagatccgg 15060 tgcagattat ttggattgag agtgaatatg agactctaat tggataccga ggggaattta 15120 tggaacgtca gtggagcatt tttgacaaga aatatttgct agctgatagt gaccttaggc 15180 gacttttgaa cgcgcaataa tggtttctga cgtatgtgct tagctcatta aactccagaa 15240 accegegget gagtggetee tteaacgttg eggttetgte agtteeaaac gtaaaacgge 15300 ttgtcccgcg tcatcggcgg gggtcataac gtgactccct taattctccg ctcatgatca 15360

WO 2004/063359 PCT/EP2004/000099 108/357

			100/33/			
gattgtcgtt	tcccgccttc	agtttaaact	atcagtgttt	gacaggatat	attggcgggt	15420
aaacctaaga	gaaaagagcg	tttattagaa	taatcggata	tttaaaaggg	cgtgaaaagg	15480
tttatccgtt	cgtccatttg	tatgtgcatg	ccaaccacag	ggttccccag	atctggcgcc	15540
ggccagcgag	acgagcaaga	ttggccgccg	cccgaaacga	tccgacagcg	cgcccagcac	15600
aggtgcgcag	gcaaattgca	ccaacgcata	cagcgccagc	agaatgccat	agtgggcggt	15660
gacgtcgttc	gagtgaacca	gatcgcgcag	gaggcccggc	agcaccggca	taatcaggcc	15720
gatgccgaca	gcgtcgagcg	cgacagtgct	cagaattacg	atcaggggta	tgttgggttt	15780
cacgtctggc	ctccggacca	gcctccgctg	gtccgattga	acgcgcggat	tctttatcac	15840
tgataagttg	gtggacatat	tatgtttatc	agtgataaag	tgtcaagcat	gacaaagttg	15900
cagccgaata	cagtgatccg	tgccgccctg	gacctgttga	acgaggtcgg	cgtagacggt	15960
ctgacgacac	gcaaactggc	ggaacggttg	ggggttcagc	agccggcgct	ttactggcac	16020
ttcaggaaca	agcgggcgct	gctcgacgca	ctggccgaag	ccatgctggc	ggagaatcat	16080
acgcattcgg	tgccgagagc	cgacgacgac	tggcgctcat	ttctgatcgg	gaatgcccgc	16140
agcttcaggc	aggcgctgct	cgcctaccgc	gatggcgcgc	gcatccatgc	cggcacgcga	16200
ccgggcgcac	cgcagatgga	aacggccgac	gcgcagcttc	gcttcctctg	cgaggcgggt	16260
ttttcggccg	gggacgccgt	caatgcgctg	atgacaatca	gctacttcac	tgttggggcc	16320
gtgcttgagg	agcaggccgg	cgacagcgat	gccggcgagc	gcggcggcac	cgttgaacag	16380
gctccgctct	cgccgctgtt	gegggeegeg	atagacgcct	tcgacgaagc	cggtccggac	16440
gcagcgttcg	agcagggact	cgcggtgatt	gtcgatggat	tggcgaaaag	gaggctcgtt	16500
gtcaggaacg	ttgaaggacc	gagaaagggt	gacgattgat	caggaccgct	gccggagcgc	16560
aacccactca	ctacagcaga	gccatgtaga	caacatcccc	teceettte	caccgcgtca	16620

WO 2004/063359 PCT/EP2004/000099 109/357

			107/357			
gacgcccgta	gcagcccgct	acgggctttt	tcatgccctg	ccctagcgtc	caagcctcac	16680
ggccgcgctc	ggcctctctg	geggeettet	ggcgctcttc	cgcttcctcg	ctcactgact	16740
cgctgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaag	gcggtaatac	16800
ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	16860
aggccaggaa	ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	cgcccccctg	16920
acgagcatca	caaaaatcga	cgctcaagtc	agaggtggcg	aaacccgaca	ggactataaa	16980
gataccaggc	gtttccccct	ggaagctccc	tegtgegete	tcctgttccg	accetgeege	17040
ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	ggcgcttttc	cgctgcataa	17100
ccctgcttcg	gggtcattat	agcgattttt	tcggtatatc	catccttttt	cgcacgatat	17160
acaggatttt	gccaaagggt	tcgtgtagac	tttccttggt	gtatccaacg	gcgtcagccg	17220
ggcaggatag	gtgaagtagg	cccacccgcg	agcgggtgtt	ccttcttcac	tgtcccttat	17280
tcgcacctgg	cggtgctcaa	cgggaatcct	gctctgcgag	gctggccggc	taccgccggc	17340
gtaacagatg	agggcaagcg	gatggctgat	gaaaccaagc	caaccaggaa	gggcagccca	17400
cctatcaagg	tgtactgcct	tccagacgaa	cgaagagcga	ttgaggaaaa	ggcggcggcg	17460
gccggcatga	gcctgtcggc	ctacctgctg	gccgtcggcc	agggctacaa	aatcacgggc	17520
gtcgtggact	atgagcacgt	ccgcgagctg	gcccgcatca	atggcgacct	gggccgcctg	17580
ggcggcctgc	tgaaactctg	gctcaccgac	gacccgcgca	cggcgcggtt	cggtgatgcc	17640
acgatecteg	ccctgctggc	gaagatcgaa	gagaagcagg	acgagcttgg	caaggtcatg	17700
atgggcgtgg	tccgcccgag	ggcagagcca	tgacttttt	agccgctaaa	acggccgggg	17760
ggtgcgcgtg	attgccaagc	acgtccccat	gcgctccatc	aagaagagcg	acttcgcgga	17820
gctggtgaag	tacatcaccg	acgagcaagg	caagaccgag	cgcctttgcg	acgctca	17877

WO 2004/063359 PCT/EP2004/000099

<210> 38 <211> 17238 <212> DNA <213> Artificial <220> <223> Plasmid <220> <221> misc_feature <222> (10264)..(10264) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10472)..(10472) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10563)..(10563) <223> n is a, c, g, or t <400> 38 cegggetggt tgccctegce getgggetgg eggeegteta tggccetgea aacgegecag 60 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 120 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180 ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240 cagcetegea aateggegaa aacgeetgat tttacgegag ttteccacag atgatgtgga 300 caagectggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360 gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 420

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgccgt

480

WO 2004/063359 PCT/EP2004/000099

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc 600 cttctcgaac cctcccggcc cgCtaacgcg ggcctcccat cccccaggg gctgcgccc 660 teggeegega aeggeeteae eecaaaaatg geagegetgg eagteettge eattgeeggg 720 atcggggcag taacgggatg ggcgatcagc ccgagcgcga cgcccggaag cattgacgtg 780 840 ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 900 ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 960 gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020 acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080 acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140 agataatata tottttatat agaagatato googtatgta aggatttoag ggggcaaggo 1200 ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260 1320 ctaatgcttg aaacccagga caataacctt atagcttgta aattctatca taattgggta 1380 atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc 1440 agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag gtgctgcctc agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt 1500 cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag 1560 ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc 1620 gcaacaaccg tetteeggag actgteatac gegtaaaaca geeagegetg gegegattta 1680 geocegacat ageoceactg tregteeatt teegegeaga egatgaegte actgeoegge 1740

WO 2004/063359 PCT/EP2004/000099 112/357

tgtatgcgcg aggttaccga ctgcgg	Cctg agtttttaa	ı gtgacgtaaa	atcgtgttga	1800
ggccaacgcc cataatgcgg gctgtt	gccc ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg gtgcgtaccg ggttga	gaag cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg agagcagaga tagcgc	tgat gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt agctgaacag gaggga	cage tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca tcatacacta aatcag	taag ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc cgtcgatact atgtta	tacg ccaactttga	aaacaacttt	gaaaaagctg	2160
ttttctggta tttaaggttt tagaat	gcaa ggaacagtga	attggagttc	gtcttgttat	2220
aattagette ttggggtate tttaaa	tact gtagaaaaga	ggaaggaaat	aataaatggc	2280
taaaatgaga atatcaccgg aattga	aaaa actgatcgaa	aaataccgct	gcgtaaaaga	2340
tacggaagga atgtctcctg ctaagg	tata taagctggtg	ggagaaaatg	aaaacctata	2400
tttaaaaatg acggacagcc ggtata	aagg gaccacctat	gatgtggaac	gggaaaagga	2460
catgatgeta tggctggaag gaaage	tgcc tgttccaaag	gtcctgcact	ttgaacggca	2520
tgatggctgg agcaatctgc tcatga	gtga ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa caaagccctg aaaaga	tat cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc gacatatcgg attgtc	ccta tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac ttactgaata acgatc	:ggc cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactecattt aaagateege gegage	gta tgattttta:	aagacggaaa	agcccgaaga	2820
ggaacttgtc ttttcccacg gcgacc	ggg agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc tttattgatc ttgggag	Jaag cggcagggcg	gacaagtggt	atgacattgc	2940

WO 2004/063359 PCT/EP2004/000099 113/357

			110,00			
cttctgcgtc	cggtcgatca	gggaggatat	cggggaagaa	cagtatgtcg	agctattttt	3000
tgacttactg	gggatcaagc	ctgattggga	gaaaataaaa	tattatattt	tactggatga	3060
attgttttag	tacctagatg	tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tcttccgcat	caagtgtttt	ggctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt	attcgtgcag	ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg	gaccgacttc	attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa	tcaggaataa	gggcacattg	cccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc	cgaggatgcc	gaaaccatcg	caagccgcac	cgtcatgcgt	gcgccccgcg	3480
aaaccttcca	gtccgtcggc	tcgatggtcc	agcaagctac	ggccaagatc	gagcgcgaca	3540
gcgtgcaact	ggctcccct	gccctgcccg	cgccatcggc	cgccgtggag	cgttcgcgtc	3600
gtctcgaaca	ggaggcggca	ggtttggcga	agtcgatgac	catcgacacg	cgaggaacta	3660
tgacgaccaa	gaagcgaaaa	accgccggcg	aggacctggc	aaaacaggtc	agcgaggcca	3720
agcaggccgc	gttgctgaaa	cacacgaagc	agcagatcaa	ggaaatgcag	ctttccttgt	3780
tcgatattgc	gccgtggccg	gacacgatgc	gagcgatgcc	aaacgacacg	gcccgctctg	3840
ccctgttcac	cacgcgcaac	aagaaaatcc	cgcgcgaggc	gctgcaaaac	aaggtcattt	3900
tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctatc	ggcgagccga	4020
tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggc	cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg	gcacctggaa	teggtgtege	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200

WO 2004/063359 PCT/EP2004/000099 114/357

			114/05/			
gcaagaaaac	gtecegttge	caggtcctga	tcgacgagga	aatcgtcgtg	ctgtttgctg	4260
gcgaccacta	cacgaaattc	atatgggaga	agtaccgcaa	gctgtcgccg	acggcccgac	4320
ggatgttcga	ctatttcagc	tcgcaccggg	agccgtaccc	gctcaagctg	gaaaccttcc	4380
gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
cctgcgaaga	gttgcgaggc	agcggcctgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	gcagccagcg	4560
ctttactggc	atttcaggaa	caagegggca	ctgctcgacg	cacttgcttc	gctcagtatc	4620
gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
tgtgattaag	gctcagattc	gacggcttgg	ageggeegae	gtgcaggatt	tccgcgagat	4740
ccgattgtcg	gccctgaaga	aagctccaga	gatgttcggg	tccgtttacg	agcacgagga	4800
gaaaaagccc	atggaggcgt	tcgctgaacg	gttgcgagat	gccgtggcat	teggegeeta	4860
catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagcgaggcc	gaggggtcgc	4980
cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	tccgacagat	5040
tccaacggga	atctggtgga	tgcgcatctt	catcctcggc	gcacttaata	tttcgctatt	5100
ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160
cgctgtgcag	ccgctgatgg	tcgtgttcat	ct¢tgecgct	ctgctaggta	gcccgatacg	5220
attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	tggtgttgac	5280
accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	cggtttccat	5340
ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	tcacctttac	5400
cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	tgtttgatcc	5460

WO 2004/063359 PCT/EP2004/000099 115/357

•			112/32/			
gċcaatcccg	atgcctacag	gaaccaatgt	tctcggcctg	gcgtggctcg	gcctgatcgg	5520
agcgggttta	acctacttcc	tttggttccg	ggggatctcg	cgactcgaac	ctacagttgt	5580
ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	tgcatcaggc	5640
cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	ataggggagt	5700
tgatatcgtc	aacgttcact	tctaaagaaa	tagcgccact	cagcttcctc	agcggcttta	5760
tccagcgatt	tcctattatg	tcggcatagt	tctcaagatc	gacagcctgt	cacggttaag	5820
cgagaaatga	ataagaaggc	tgataattcg	gatctctgcg	agggagatga	tatttgatca	5880
caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	gatcatccgt	5940
gtttcaaacc	cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	atgagcaaag	6000
tctgccgcct	tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	ctgcctgtat	6060
cgagtggtga	ttttgtgccg	agctgccggt	cggggagctg	ttggctggct	ggtggcagga	6120
tatattgtgg	tgtaaacaaa	ttgacgctta	gacaacttaa	taacacattg	cggacgtttt	6180
taatgtactg	gggtggtttt	tcttttcacc	agtgagacgg	gcaacagctg	attgcccttc	6240
accgcctggc	cctgagagag	ttgcagcaag	cggtccacgc	tggtttgccc	cagcaggcga	6300
aaatcctgtt	tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	6360
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480
cacccaaatc	aagtttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	6540
ggageceeeg	atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600
agaaagcgaa	aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt	cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720

WO 2004/063359 PCT/EP2004/000099 116/357

tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag 6780 ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840 cgaggagect cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6900 gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcatc tgtacttcaa 6960 gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7020 aagatecaeg tettegggaa geeagegaet ggtgaeetee agegteeett taaggetgee 7080 aacagctttc tcagccaggg ccagcccaag accgacaagg cctccctcca gaacgccgag 7140 aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200 cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg 7260 tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg .7320 teetttgget tteaatatte teaagtatet egagtttgaa ettatteet gtgaacettt 7380 tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7440 tegaaataca teeggatgte gaaggettgg ggeaectgeg ttggttgaat ttagaacgtg 7500 gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7560 tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga 7620 tetetagtta geteaceaca aaagteagae ggegtaacea aaagteacae aacacaaget 7680 gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740 atttaattet atttgtgttt gategagaee taatacagee eetacaaega eeatcaaagt 7800 cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860 ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatc 7920 tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980

cctagctgat	tctggagtga	cccagagggt	catgacttga	gcctaaaatc	cgccgcctcc	8040
accatttgta	gaaaaatgtg	·acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100
tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataagc	cactggccag	8160
acagctctgg	cggctctgag	gtgcagtgga	tgattattaa	tccgggaccg	gccgcccctc	8220
cgccccgaag	tggaaaggct	ggtgtgcccc	tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatagccgt	cggcgaaata	gcatgccatt	aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	agattcacga	gatagtacct	tctccgaagt	aggtagagcg	agtacccggc	8460
gcgtaagctc	cctaattggc	ccatccggca	tctgtagggc	gtccaaatat	cgtgcctctc	8520
ctgctttgcc	cggtgtatga	aaccggaaag	gccgctcagg	agctggccag	cggcgcagac	8580
cgggaacaca	agctggcagt	cgacccatcc	ggtgctctgc	actcgacctg	ctgaggtccc	8640
tcagtccctg	gtaggcagct	ttgccccgtc	tgtccgcccg	gtgtgtcggc	ggggttgaca	8700
aggtcgttgc	gtcagtccaa	catttgttgc	catattttcc	tgetetecee	accagctgct	8760
cttttcttt	ctctttcttt	teccatette	agtatattca	tcttcccatc	caagaacctt	8820
tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240

WO 2004/063359 PCT/EP2004/000099 118/357

tgtcacgttg caagacctgc ctgaaaccga actgeccgct gttctgcag	gc cggtcgcgga	9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggtto	g gcccattcgg	9360
accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcg	ga ttgctgatcc	9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtco	g tegegeagge	9480
tetegatgag etgatgettt gggeegagga etgeeeegaa gteeggead	c tegtgeaege	9540
ggatttegge tecaacaatg teetgaegga caatggeege ataacageg	gg tcattgactg	9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttc	t tctggaggcc	9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcat	c cggagcttgc	9720
aggategeeg eggeteeggg egtatatget eegeattggt ettgaceaa	c tctatcagag	9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgc	g acgcaatcgt	9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagc	g cggccgtctg	9900
gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccc	a gcactcgtcc	9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaa	c gttactgaaa	10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgt	c agctccggag	10080
ttgagacaaa tggtgttcag gatctcgata agatacgttc atttgtcca	a gcagcaaaga	10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgt	t ttcgggttta	10200
cctcttccag atacagetca tetgcaatge attaatgeat tgactgeaa	c ctagtaacgc	10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcat	t ttcgggagac	10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc	c cgctcttggc	10380
tocacgogae tatatatttg tototaattg tactttgaca tgctcctct	ctttactctg	10440
atagettgae tatgaaaatt eegteaceag eneetgggtt egeaaagata	a attgcatgtt	10500

WO 2004/063359 PCT/EP2004/000099 119/357

tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
canttectae	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactete	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tctccactcg	acctgcaggc	atgcaagctt	ctaccgcttg	10800
gaccagtcca	gttccaggac	cagtcgctcc	acctcctccg	ccgcacctgg	aatgtgctgc	10860
agctcctgtg	gacccaagaa	cataccccag	ggcgcgccac	cgtacttgcc	gctgtggtgt	10920
agctggtggg	ccactgtcag	gcgcttcatg	tagggcaggc	cagcgatggg	cccggtggga	10980
aagcgcctgt	gcaccaggcc	atcgtgtaca	aacatatatg	ccatgccgta	tagcgtgatg	11040
cccagccccg	ctccaaagca	ggccgccccc	aggacgttgg	gcagccagaa	gccaaaggta	11100
cacaggagca	tggcgggcag	tccattgatg	attgcaaaca	agtcgttggc	ttcaaagggt	11160
ccagtgcgag	gtgtgtggtg	gctcttgtgc	agcagccagc	ccagaggcga	ctcatgccag	11220
atggctttgt	gtgcatagcg	ggcatacatc	tccatgccga	gcgcgccacc	aaccaccaag	11280
aggagagtgc	cagccacttc	accccatggc	actgegeege	ccacggtcat	gtgcatggca	11340
aatctcaggt	aggtggcgaa	gatggcaatg	cctgacacgc	caattgatgc	tgcaatggcg	11400
gcagcctggt	atgacagctg	ctcccgtttg	cgccgggcac	gacgctctgc	gatagcccgg	11460
tcaagctgct	ggagtgctac	ateggegetg	tgctcatcgc	ccgcgccggc	agcctgcacg	11520
gttcccagcg	cctcctctgt	ctgtggtgct	gccactcgca	gccgaactaa	cgagcaccgc	11580
tgagcatgca	ggcagacttt	gggccgcgtg	atgtcgcggg	ctagttcaac	gcggcgggcc	11640
ttgacgctga	ttgactgcag	cttcgacagc	atagagataa	aataaaaaga	gaagaaaaga	11700
aagtttgtac	aatttcttt	tgtttatata	acatacacgc	tatgtcaaca	tttagaataa	11760

WO 2004/063359 PCT/EP2004/000099 120/357

gggggaaaaa	atcttccatc	atattcgaat	gcacaagatt	atttctttgt	tegetetttt	11820
tggtcgggtc	atcgagattt	agagtgtaat	caaagatact	gtcatctcga	gagcgttgca	11880
caggctgctg	tttgccaaat	tggatgtttg	ccgaattagt	aaaatacgca	agcatttctt	11940-
acctttccgc	tecettttee	taattctccc	aaagactaaa	tgaggaaaga	taaaggacaa	12000
agaaaatgta	aagacaaaga	aattgaaaac	gatataaact	tgcagcacgt	aagaccaaag	12060
caaattggta	actattcttg	tgtacaaaca	tgtataaaaa	aaaacttttt	tttgctcctg	12120
gaggacaaaa	tttcaaactc	cțtgaagaag	attgcttgta	tatctatcat	atgcatatat	12180
catatcgatg	gaaaaagaaa	gtcaggcatg	tatttataaa	aagaagaatg	tgccatgctt	12240
ccgaatttct	tttcactttc	ttttccttat	ctattttaat	ctcaagcttg	gcgtaatcat	12300
ggtcatagct	gtttcctgtg	tgaaattgtt	atccgctcac	aattccacac	aacatacgag	12360
ccggaagcat	aaagtgtaaa	gcctggggtg	cctaatgagt	gagctaactc	acattaattg	12420
cgttgcgctc	actgcccgct	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	12480
tcggccaacg	cgcggggaga	ggcggtttgc	gtattgggcc	aaagacaaaa	gggcgacatt	12540
caaccgattg	agggagggaa	ggtaaatatt	gacggaaatt	attcattaaa	ggtgaattat	12600
caccgtcacc	gacttgagcc	atttgggaat	tagagccagc	aaaatcacca	gtagcaccat	12660
taccattagc	aaggccggaa	acgtcaccaa	tgaaaccatc	gatagcagca	ccgtaatcag	12720
tagcgacaga	atcaagtttg	cctttagcgt	cagactgtag	cgcgttttca	tcggcatttt	12780
cggtcatagc	ccccttatta	gcgtttgcca	tcttttcata	atcaaaatca	ccggaaccag	12840
agccaccacc	ggaaccgcct	ccctcagagc	cgccaccctc	agaaccgcca	ccctcagagc	12900
caccaccctc	agagccgcca	ccagaaccac	caccagagcc	gccgccagca	ttgacaggag	12960
gcccgatcta	gtaacataga	tgacaccgcg	cgcgataatt	tatcctagtt	tgcgcgctat	13020

WO 2004/063359 PCT/EP2004/000099 121/357

attttgtttt	ctatcgcgta	ttaaatgtat	aattgcggga	ctctaatcat	aaaaacccat	13080
ctcataaata	acgtcatgca	ttacatgtta	attattacat	gcttaacgta	attcaacaga	13140
aattatatga	taatcatcgc	aagaccggca	acaggattca	atcttaagaa	actttattgc	13200
caaatgtttg	aacgatcggg	gatcatccgg	gtctgtggcg	ggaactccac	gaaaatatcc	13260
gaacgcagca	agatatcgcg	gtgcatctcg	gtcttgcctg	ggcagtcgcc	gccgacgccg	13320
ttgatgtgga	cgccgggccc	gatcatattg	tcgctcagga	tcgtggcgtt	gtgcttgtcg	13380
gccgttgctg	tcgtaatgat	atcggcacct	tcgaccgcct	gttccgcaga	gatcccgtgg	13440
gcgaagaact	ccagcatgag	atccccgcgc	tggaggatca	tecageegge	gtcccggaaa	13500
acgattccga	agcccaacct	ttcatagaag	gcggcggtgg	aatcgaaatc	tcgtgatggc	13560
aggttgggcg	tegettggte	ggtcatttcg	aaccccagag	tecegeteag	aagaactcgt	13620
caagaaggcg	atagaaggcg	atgcgctgcg	aatcgggagc	ggcgataccg	taaagcacga	13680
ggaagcggtc	agcccattcg	ccgccaagct	cttcagcaat	atcacgggta	gccaacgcta	13740
tgtcctgata	geggteegee	acacccagcc	ggccacagtc	gatgaatcca	gaaaagcggc	13800
cattttccac	catgatattc	ggcaagcagg	catcgccatg	ggtcacgacg	agatcatcgc	13860
cgtcgggcat	gcgcgccttg	agcctggcga	acagttcggc	tggcgcgagc	ccctgatgct	13920
cttcgtccag	atcatcctga	tcgacaagac	cggcttccat	ccgagtacgt	gctcgctcga	13980
tgcgatgttt	cgcttggtgg	tcgaatgggc	aggtagccgg	atcaagcgta	tgcagccgcc	14040
gcattgcatc	agccatgatg	gatactttct	cggcaggagc	aaggtgagat	gacaggagat	14100
cctgccccgg	cacttcgccc	aatagcagcc	agtcccttcc	cgcttcagtg	acaacgtcga	14160
gcacagetge	gcaaggaacg	cccgtcgtgg	ccagccacga	tageegeget	gcctcgtcct	14220
gcagttcatt	cagggcaccg	gacaggtcgg	tcttgacaaa	aagaaccggg	cgcccctgcg	14280

WO 2004/063359 PCT/EP2004/000099

ctgacagccg gaacacggcg gcatcagagc agccgattgt ctgttgtgcc cagtcatagc 14340 cgaatageet etecacecaa geggeeggag aacetgegtg caatecatet tgttcaatea 14400 tgcgaaacga tccagatccg gtgcagatta tttggattga gagtgaatat gagactctaa 14460 ttggataccg aggggaattt atggaacgtc agtggagcat ttttgacaag aaatatttgc 14520 tagctgatag tgaccttagg cgacttttga acgcgcaata atggtttctg acgtatgtgc 14580 ttagctcatt aaactccaga aacccgcggc tgagtggctc cttcaacgtt gcggttctgt 14640 cagttccaaa cgtaaaacgg cttgtcccgc gtcatcggcg ggggtcataa cgtgactccc 14700 ttaattctcc gctcatgatc agattgtcgt ttcccgcctt cagtttaaac tatcagtgtt 14760 tgacaggata tattggcggg taaacctaag agaaaagagc gtttattaga ataatcggat 14820 atttaaaagg gcgtgaaaag gtttatccgt tcgtccattt gtatgtgcat gccaaccaca 14880 gggttcccca gatctggcgc cggccagcga gacgagcaag attggccgcc gcccgaaacg 14940 atccgacagc gcgcccagca caggtgcgca ggcaaattgc accaacgcat acagcgccag 15000 cagaatgcca tagtgggcgg tgacgtcgtt cgagtgaacc agatcgcgca ggaggcccgg 15060 cagcaccggc ataatcaggc cgatgccgac agcgtcgagc gcgacagtgc tcagaattac 15120 gatcaggggt atgttgggtt tcacgtctgg cctccggacc agcctccgct ggtccgattg 15180 aacgcgcgga ttctttatca ctgataagtt ggtggacata ttatgtttat cagtgataaa 15240 gtgtcaagca tgacaaagtt gcagccgaat acagtgatcc gtgccgccct ggacctgttg 15300 aacgaggtcg gcgtagacgg tctgacgaca cgcaaactgg cggaacggtt gggggttcag 15360 cagccggcgc tttactggca cttcaggaac aagcgggcgc tgctcgacgc actggccgaa 15420 gccatgctgg cggagaatca tacgcattcg gtgccgagag ccgacgacga ctggcgctca 15480 tttctgatcg ggaatgcccg cagcttcagg caggcgctgc tcgcctaccg cgatggcgcg 15540

WO 2004/063359 PCT/EP2004/000099 123/357

		•				
cgcatccatg	ccggcacgcg	accgggcgca	ccgcagatgg	aaacggccga	cgcgcagctt	15600
egetteetet	gcgaggcggg	tttttcggcc	ggggacgccg	tcaatgcgct	gatgacaatc	15660
agctacttca	ctgttggggc	cgtgcttgag	gagcaggccg	gcgacagcga	tgccggcgag	15720
cgcggcggca	ccgttgaaca	ggeteegete	tegeegetgt	tgcgggccgc	gatagacgcc	15780
ttcgacgaag	ccggtccgga	cgcagcgttc	gagcagggac	tcgcggtgat	tgtcgatgga	15840
ttggcgaaaa	ggaggctcgt	tgtcaggaac	gttgaaggac	cgagaaaggg	tgacgattga	15900
tcaggaccgc	tgccggagcg	caacccactc	actacagcag	agccatgtag	acaacatccc	15960
ctccccttt	ccaccgcgtc	agacgcccgt	agcagcccgc	tacgggcttt	ttcatgccct	16020
gccctagcgt	ccaagcctca	eggeegeget	cggcctctct	ggcggccttc	tggcgctctt	16080
ccgcttcctc	gctcactgac	tegetgeget	cggtcgttcg	gctgcggcga	gcggtatcag	16140
ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	16200
tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcġtttt	16260
tccataggct	ccgccccct	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	16320
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaagctcc	ctcgtgcgct	16380
ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tegggaageg	16440
tggcgctttt	ccgctgcata	accetgette	ggggtcatta	tagcgatttt	ttcggtatat	16500
ccatcctttt	tcgcacgata	tacaggatt	tgccaaaggg	ttcgtgtaga	ctttccttgg	16560
tgtatccaac	ggcgtcagco	gggcaggata	. ggtgaagtag	geceaceege	: gagcgggtgt	16620
tccttcttca	ı ctgtccctta	ttegeacetg	gcggtgctca	acgggaatco	: tgctctgcga	16680
ggctggccgg	g ctaccgccgg	g cgtaacagat	gagggcaagc	ggatggctga	a tgaaaccaag	16740
ccaaccagga	a agggcagcco	e acctatcaag	gtgtactgco	: ttccagacga	a acgaagagcg	16800

attgaggaaa aggeggegge ggeeggeatg ageetgtegg cetacetget ggeeggee 16860

cagggetaca aaatcacggg egtegtggae tatgagcaeg teeggaget ggeeegcate 16920

aatggegace tgggeegeet gggeggeetg etgaaactet ggeteacega egaceeggee 16980

acggegggt teeggtgatge cacgateete geeetgetgg egaagatega agagaageag 17040

gaegagettg geaaggteat gatggeggt gteegeeega gggeagagee atgaetttt 17100

tageegetaa aacggeeggg gggtgeget gattgeeaag eacgteecea tgegeteeat 17160

caagaagage gaettegegg agetggtaa gtacateace gaegageaag geaagaeega 17220

gegeetttge gaegetea

<210> 39

<211> 17238

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

WO 2004/063359 PCT/EP2004/000099 125/357

<400> 39 ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 120 180 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 240 ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 300 caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360 gagggggggg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 420 tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 540 600 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgccccc cttctcgaac cctcccggcc cgctaacgcg ggcctcccat cccccaggg gctgcgccc 660 tcggccgcga acggcctcac cccaaaaatg gcagcgctgg cagtccttgc cattgccggg 720 780 atcggggcag taacgggatg ggcgatcagc ccgagcgcga cgcccggaag cattgacgtg ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840 ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900 gcaattttta ccttgggcat tcttggcata gtggtcgcgg gtgccgtgct cgtgttcggg 960 1020 ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080 acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140 1200 agataatata tettttatat agaagatate geegtatgta aggattteag ggggeaagge ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

WO 2004/063359 PCT/EP2004/000099 126/357

•						
ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgeegete	aattcgctgc	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg	tcttccggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat	agccccactg	ttcgtccatt	teegegeaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg	aggttaccga	ctgcggcctg	agttttttaa	gtgacgtaaa	atcgtgttga	1800
ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg	agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc	cgtcgatact	atgttatacg	ccaactttga	aaacaacttt	: gaaaaagctg	2160
ttttctggta	tttaaggttt	tagaatgcaa	ı ggaacagtga	attggagtto	gtcttgttat	2220
aattagcttc	ttggggtato	: tttaaatact	gtagaaaaga	a ggaaggaaat	aataaatggc	2280
taaaatgaga	atatcaccgg	g aattgaaaaa	a actgatcgaa	aaataccgct	: gcgtaaaaga	2340
tacggaagga	atgtctcctg	g ctaaggtata	a taagetggtg	g ggagaaaat	g aaaacctata	2400
tttaaaaatg	acggacagco	ggtataaagg	g gaccacctat	gatgtggaa	c gggaaaagga	2460
catgatgcta	tggctggaag	g gaaagctgc	c tgttccaaag	g gtcctgcac	t ttgaacggca	2520

WO 2004/063359 PCT/EP2004/000099 127/357

tgatggctgg agcaatc	tgc tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa caaagcc	ctg aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc gacatat	cgg attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac ttactga	ata acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt aaagatc	cgc gcgagctgta	tgatttttta	aagacggaaa	agcccgaaga	2820
ggaacttgtc ttttccc	acg gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc tttattg	atc ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc cggtcga	tca gggaggatat	cggggaagaa	cagtatgtcg	agctattttt	3000
tgacttactg gggatca	agc ctgattggga	gaaaataaaa	tattatattt	tactggatga	3060
attgttttag tacctag	atg tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tcttccgcat caagtgt	ttt ggctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt attcgtg	cag ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg gaccgac	ttc attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa tcaggaa	itaa gggcacattg	ccccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat gaatcgg	gacg tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc cgagga	gcc gaaaccatcg	caagccgcac	cgtcatgcgt	gegeeeegeg	3480
aaaccttcca gtccgt	egge tegatggtoc	agcaagctac	ggccaagato	gagcgcgaca	3540
gegtgeaact ggetee	cect geoctgooog	cgccatcggc	: cgccgtggag	cgttcgcgtc	3600
gtctcgaaca ggaggc	ggca ggtttggcga	agtcgatgac	categacaeg	g cgaggaacta	3660
tgacgaccaa gaagcg	aaaa accgccggcg	g aggaeetgge	e aaaacaggto	agcgaggcca	3720
agcaggccgc gttgct	gaaa cacacgaago	agcagatcaa	a ggaaatgcag	g ctttccttgt	3780

WO 2004/063359 PCT/EP2004/000099 128/357

togatattgo googtggoog gacacgatgo gagogatgoo aaacgacaog goocgototg	3840
ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt	3900
tccacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg	3960
acgaactggt gtggcagcag gtgttggagt acgcgaagcg cacccctatc ggcgagccga	4020
tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt	4080
acacgaagge egaggaatge etgtegegee tacaggegae ggegatggge tteaegteeg	4140
accgcgttgg gcacctggaa tcggtgtcgc tgctgcaccg cttccgcgtc ctggaccgtg	4200
gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg	4260
gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac	4320
ggatgttcga ctatttcage tcgcaccggg agccgtaccc gctcaagctg gaaaccttcc	4380
gcctcatgtg cggatcggat tccacccgcg tgaagaagtg gcgcgagcag gtcggcgaag	4440
cctgcgaaga gttgcgaggc agcggcctgg tggaacacgc ctgggtcaat gatgacctgg	4500
tgcattgcaa acgctagggc cttgtggggt cagttccggc tgggggttca gcagccagcg	4560
ctttactggc atttcaggaa caagcgggca ctgctcgacg cacttgcttc gctcagtatc	4620
gctcgggacg cacggcgcgc tctacgaact gccgataaac agaggattaa aattgacaat	4680
tgtgattaag gctcagattc gacggcttgg agcggccgac gtgcaggatt tccgcgagat	4740
ccgattgtcg gccctgaaga aagctccaga gatgttcggg tccgtttacg agcacgagga	4800 [°]
gaaaaagccc atggaggcgt tcgctgaacg gttgcgagat gccgtggcat tcggcgccta	4860
catcgacggc gagatcattg ggctgtcggt cttcaaacag gaggacggcc ccaaggacgc	4920
tcacaaggcg catctgtccg gcgttttcgt ggagcccgaa cagcgaggcc gaggggtcgc	4980
cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat	5040

tccaacggga	atctggtgga	tgcgcatctt	catcctcggc	gcacttaata	tttcgctatt	5100
ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160
cgctgtgcag	ccgctgatgg	tcgtgttcat	ctctgccgct	ctgctaggta	gcccgatacg	5220
attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	tggtgttgac	5280
accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	cggtttccat	5340
ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	tcacctttac	5400
cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	tgtttgatcc	5460
gccaatcccg	atgcctacag	gaaccaatgt	tctcggcctg	gcgtggctcg	gcctgatcgg	5520
agcgggttta	acctacttcc	tttggttccg	ggggatctcg	cgactcgaac	ctacagttgt	5580
ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	tgcatcaggc	5640
cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	ataggggagt	5700
tgatatcgtc	aacgttcact	tctaaagaaa	tagcgccact	cagcttcctc	agcggcttta	5760
tccagcgatt	tcctattatg	tcggcatagt	tctcaagatc	gacagectgt	cacggttaag	5820
cgagaaatga	ataagaaggc	tgataattcg	gatctctgcg	agggagatga	tatttgatca	5880
caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	gatcatccgt	5940
gtttcaaacc	: cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	atgagcaaag	6000
tctgccgcct	: tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	g ctgcctgtat	6060
cgagtggtga	ı ttttgtgccg	agctgccggt	: cggggagctg	g ttggctggct	ggtggcagga	6120
tatattgtgg	y tgtaaacaaa	ttgacgctta	gacaacttaa	a taacacatto	g cggacgtttt	6180
taatgtact	g gggtggtttt	tcttttcacc	agtgagacgg	g gcaacagcts	g attgcccttc	6240
accgcctgg	c cctgagagag	g ttgcagcaag	g cggtccacgo	c tggtttgcc	c cagcaggcga	6300

WO 2004/063359 PCT/EP2004/000099 130/357

1 1 . 1 . 1		h		ttataaatca	222cantago	6360	
aaatcctgtt	tgatggtggt	teegaaateg	gcaaaatccc	LLaLaaacca	aaagaacagc	0300	
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420	
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480	
cacccaaatc	aagtttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aacccțaaag	6540	
ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600 .	
agaaagcgaa	aggageggge	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660	
cgggcctctt	cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720	
tgggtaacgc	cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780	
ctcggtaccc	ggggatcttt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840	
cgaggagcct	cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900	
gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960	
gtacaccaac	gctcttctga	tecagtegat	catccgctga	aggcgctttc	gaatctggtt	7020	
aagatccacg	tcttcgggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080	
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140	
aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200	
cggtgtcaag	aggatattct	tegaetetgt	attatagata	agatgatgag	gaattggagg	7260	
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320	
tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380	
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440	
tcgaaataca	tccggatgtc	gaaggettgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500	
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560	

WO 2004/063359 PCT/EP2004/000099 131/357

tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaạcatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ggtggagagc	ttataccgag	ctcccaaatc	7920
tgtccagatc	atggttgacc	ggtgcctgga	tcttcctata	gaatcatcct	tattcgttga	7980 ·
cctagctgat	tctggagtga	cccagagggt	catgacttga	gcctaaaatc	cgccgcctcc	8040
accatttgta	gaaaaatgtg	acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100
tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataagc	cactggccag	8160
acagctctgg	cggctctgag	gtgcagtgga	tgattattaa	tccgggaccg	geegeeeete	8220
cgccccgaag	tggaaaggct	ggtgtgcccc	tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatagccgt	cggcgaaata	gcatgccatt	aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	agattcacga	gatagtacct	tctccgaagt	aggtagagcg	agtacccggc	8460
gcgtaagctc	cctaattggc	ccatccggca	tctgtagggc	gtccaaatat	cgtgcctctc	8520
ctgctttgcc	: cggtgtatga	aaccggaaag	gccgctcagg	agctggccag	cggcgcagac	8580
cgggaacaca	agctggcagt	cgacccatco	ggtgctctgc	actcgacctg	ctgaggtccc	8640
tcagtccctg	gtaggcagct	. ttgccccgtc	tgtccgcccg	gtgtgtcggc	ggggttgaca	8700
aggtcgttgc	gtcagtccaa	catttgttgc	catatttcc	tgctctcccc	: accágctgct	8760
cttttcttt	ctctttctt	: tcccatcttc	: agtatattca	tcttcccatc	: caagaacett	8820

WO 2004/063359 PCT/EP2004/000099

tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tectteccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagettegat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	gcccattcgg	9360
accgcaagga	atcggtcaat	acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	tegegeagge	9480
tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	tcgtgcacgc	9540
ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	cggagcttgc	9720
aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga	gccgggactg	tegggegtae	acaaatcgcc	cgcagaagcg	cggccgtctg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960
gagggcaaag	gaatagagta	gatgccgacc	gcgggatcga	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080

WO 2004/063359 PCT/EP2004/000099 133/357

ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttcgggttta	10200
cctcttccag	atacagctca	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagagc	tattttcatt	ttcgggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380
tccacgcgac	tatatatttg	tctctaattg	tactttgaca	tgctcctctt	ctttactctg	10440
atagcttgac	tatgaaaatt	ccgtcaccag	cncctgggtt	cgcaaagata	attgcatgtt	10500
tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
cantitcctac	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactctc	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tctccactcg	acctgcaggc	atgcaagctt	agagataaaa	10800
taaaaagaga	agaaaagaaa	gtttgtacaa	tttctttttg	tttatataac	atacacgcta	10860
tgtcaacatt	tagaataagg	gggaaaaaat	cttccatcat	attcgaatgc	acaagattat	10920
ttctttgttc	gctctttttg	gtcgggtcat	cgagatttag	agtgtaatca	aagatactgt	10980
catctcgaga	gcgttgcaca	ggctgctgtt	tgccaaattg	gatgtttgcc	gaattagtaa	11040
aatacgcaag	catttcttac	ctttccgctc	ccttttccta	attctcccaa	agactaaatg	11100
aggaaagata	aaggacaaag	aaaatgtaaa	gacaaagaaa	ttgaaaacga	tataaacttg	11160
cagcacgtaa	gaccaaagca	aattggtaac	tattcttgtg	tacaaacatg	tataaaaaaa	11220
aactttttt	tgctcctgga	ggacaaaatt	tcaaactcct	tgaagaagat	tgcttgtata	11280
tctatcatat	: gcatatatca	tatcgatgga	aaaagaaagt	caggcatgta	tttataaaaa	11340

WO 2004/063359 PCT/EP2004/000099 134/357

			•			
gaagaatgtg	ccatgcttcc	gaatttcttt	tcactttctt	ttccttatct	attttaatct	11400
catgctgtcg	aagctgcagt	caatcagcgt	caaggcccgc	cgcgttgaac	tagcccgcga	11460
catcacgcgg	cccaaagtct	gcctgcatgc	tcagcggtgc	tcgttagttc	ggctgcgagt	11520
ggcagcacca	cagacagagg	aggcgctggg	aaccgtgcag	gctgccggcg	cgggcgatga	11580
gcacagcgcc	gatgtagcac	tccagcagct	tgaccgggct	atcgcagagc	gtcgtgcccg	11640
gcgcaaacgg	gagcagctgt	cataccaggc	tgccgccatt	gcagcatcaa	ttggcgtgtc	11700
aggcattgcc	atcttcgcca	cctacctgag	atttgccatg	cacatgaccg	tgggcggcgc	11760
agtgccatgg	ggtgaagtgg	ctggcactct	cctcttggtg	gttggtggcg	cgctcggcat	11820
ggagatgtat	gcccgctatg	cacacaaagc	catctggcat	gagtcgcctc	tgggctggct	11880
gctgcacaag	agccaccaca	cacctcgcac	tggaccettt	gaagccaacg	acttgtttgc	11940
aatcatcaat	ggactgcccg	ccatgctcct	gtgtaccttt	ggcttctggc	tgcccaacgt	12000
cctgggggcg	gcctgctttg	gagcggggct	gggcatcacg	ctatacggca	tggcatatat	12060
gtttgtacac	gatggcctgg	tgcacaggcg	ctttcccacc	gggcccatcg	ctggcctgcc	12120
ctacatgaag	cgcctgacag	tggcccacca	gctacaccac	agcggcaagt	acggtggcgc	12180
gccctggggt	atgttcttgg	gtccacagga	gctgcagcac	attccaggtg	, cggcggagga	12240
ggtggagcga	ctggtcctgg	aactggactg	gtccaagcgg	tagaagcttg	gcgtaatcat	12300
ggtcatagct	gtttcctgtg	tgaaattgtt	atccgctcac	aattccacac	aacatacgag	12360
ccggaagcat	aaagtgtaaa	gcctggggtg	cctaatgagt	gagctaactc	acattaattg	12420
cgttgcgctc	actgcccgct	ttccagtcgg	gaaacctgtc	gtgccagctg	cattaatgaa	12480
tcggccaacg	cgcggggaga	ggcggtttgc	gtattgggcc	aaagacaaaa	gggcgacatt	12540
caaccgattg	agggagggaa	ggtaaatatt	gacggaaatt	attcattaaa	ggtgaattat	12600

WO 2004/063359 PCT/EP2004/000099 135/357

caccgtcacc	gacttgagcc	atttgggaat	tagagccagc	aaaatcacca	gtagcaccat	12660
taccattagc	aaggccggaa	acgtcaccaa	tgaaaccatc	gatagcagca	ccgtaatcag	12720
tagcgacaga	atcaagtttg	cctttagcgt	cagactgtag	cgcgttttca	tcggcatttt	12780
cggtcatagc	ccccttatta	gcgtttgcca	tcttttcata	atcaaaatca	ccggaaccag	12840
agccaccacc	ggaaccgcct	ccctcagagc	cgccaccctc	agaaccgcca	ccctcagagc	12900
caccaccctc	agagccgcca	ccagaaccac	caccagagcc	gccgccagca	ttgacaggag	12960
gcccgatcta	gtaacataga	tgacaccgcg	cgcgataatt	tatcctagtt	tgcgcgctat	13020
attttgtttt	ctatcgcgta	ttaaatgtat	aattgcggga	ctctaatcat	aaaaacccat	13080
ctcataaata	acgtcatgca	ttacatgtta	attattacat	gcttaacgta	attcaacaga	13140
aattatatga	taatcatcgc	aagaccggca	acaggattca	atcttaagaa	actttattgc	13200
caaatgtttg	aacgatcggg	gatcatccgg	gtctgtggcg	ggaactccac	gaaaatatcc	13260
gaacgcagca	agatatcgcg	gtgcatctcg	gtcttgcctg	ggcagtcgcc	gccgacgccg	13320
ttgatgtgga	cgccgggccc	gatcatattg	tcgctcagga	tegtggegtt	gtgcttgtcg	13380
gccgttgctg	tcgtaatgat	atcggcacct	tcgaccgcct	gttccgcaga	gatecegtgg	13440
gcgaagaact	ccagcatgag	ateceegege	tggaggatca	tccagccggc	gtcccggaaa	13500
acgattccga	agcccaacct	ttcatagaag	gcggcggtgg	aatcgaaatc	tegtgatgge	13560
aggttgggcg	tcgcttggtc	ggtcatttcg	aaccccagag	tcccgctcag	aagaactcgt	13620
caagaaggcg	g atagaaggcg	atgcgctgcg	aatcgggagc	ggcgataccg	taaagcacga	13680
ggaageggte	agcccattcg	ccgccaagct	cttcagcaat	atcacgggta	gccaacgcta	13740
tgtcctgata	a geggteeged	: acacccagcc	ggccacagtc	: gatgaatcca	gaaaagcggc	13800
cattttcca	c catgatatto	ggcaagcagg	catcgccatg	ggtcacgacg	g agateatege	13860

WO 2004/063359 PCT/EP2004/000099 136/357

cgtcgggcat	gcgcgccttg	agcctggcga	acagttcggc	tggcgcgagc	ccctgatgct	13920
cttcgtccag	atcatcctga	tcgacaagac	cggcttccat	ccgagtacgt	gctcgctcga	13980
tgcgatgttt	cgcttggtgg	tcgaatgggc	aggtagccgg	atcaagcgta	tgcagccgcc	14040
gcattgcatc	agccatgatg	gatactttct	cggcaggagc	aaggtgagat	gacaggagat	14100
cctgccccgg	cacttcgccc	aatagcagcc	agtcccttcc	cgcttcagtg	acaacgtcga	14160
gcacagctgc	gcaaggaacg	cccgtcgtgg	ccagccacga	tagccgcgct	gcctcgtcct	14220
gcagttcatt	cagggcaccg	gacaggtcgg	tcttgacaaa	aagaaccggg	egecectgeg	14280
ctgacagccg	gaacacggcg	gcatcagagc	agccgattgt	ctgttgtgcc	cagtcatagc	14340
cgaatagcct	ctccacccaa	gcggccggag	aacctgcgtg	caatccatct	tgttcaatca	14400
tgcgaaacga	tccagatccg	gtgcagatta	tttggattga	gagtgaatat	gagactctaa	14460
ttggataccg	aggggaattt	atggaacgtc	agtggagcat	ttttgacaag	aaatatttgc	14520
tagctgatag	tgaccttagg	cgacttttga	acgcgcaata	atggtttctg	acgtatgtgc	14580
ttagctcatt	aaactccaga	aacccgcggc	tgagtggctc	cttcaacgtt	gcggttctgt	14640
cagttccaaa	cgtaaaacgg	cttgtcccgc	gtcatcggcg	ggggtcataa	cgtgactccc	14700
ttaattctcc	gctcatgatc	agattgtcgt	ttcccgcctt	cagtttaaac	tatcagtgtt	14760
tgacaggata	. tattggcggg	taaacctaag	agaaaagagc	gtttattaga	ataatcggat	14820
atttaaaagg	gcgtgaaaag	gtttatccgt	tegtecattt	gtatgtgcat	gccaaccaca	14880
gggttcccca	gatctggcgc	cggccagcga	gacgagcaag	attggccgcc	gcccgaaacg	14940
atccgacago	: gcgcccagca	caggtgcgca	ggcaaattgc	accaacgcat	acagcgccag	15000
cagaatgcca	a tagtgggcgg	tgacgtcgtt	cgagtgaacc	agatcgcgca	ggaggcccgg	15060
cagcaccggc	ataatcaggo	: cgatgccgac	agcgtcgagc	gcgacagtgc	: tcagaattac	15120

WO 2004/063359 PCT/EP2004/000099 137/357

gatcaggggt	atgttgggtt	tcacgtctgg	cctccggacc	agcctccgct	ggtccgattg	15180
aacgcgcgga	ttctttatca	ctgataagtt	ggtggacata	ttatgtttat	cagtgataaa	15240
gtgtcaagca	tgacaaagtt	gcagccgaat	acagtgatcc	gtgccgccct	ggacctgttg	15300
aacgaggtcg	gcgtagacgg	tctgacgaca	cgcaaactgg	cggaacggtt	gggggttcag	15360
cageeggege	tttactggca	cttcaggaac	aagcgggcgc	tgctcgacgc	actggccgaa	15420
gccatgctgg	cggagaatca	tacgcattcg	gtgccgagag	ccgacgacga	ctggcgctca	15480
tttctgatcg	ggaatgcccg	cagcttcagg	caggegetge	tegectaceg	cgatggcgcg	15540
cgcatccatg	ceggcaegeg	accgggcgca	ccgcagatgg	aaacggccga	cgcgcagctt	15600
cgcttcctct	gcgaggcggg	ttttcggcc	ggggacgccg	tcaatgcgct	gatgacaatc	15660
agctacttca	ctgttggggc	cgtgcttgag	gagcaggccg	gcgacagcga	tgccggcgag	15720
cgcggcggca	ccgttgaaca	ggctccgctc	tegeegetgt	tgcgggccgc	gatagacgcc	15780
ttcgacgaag	ccggtccgga	cgcagcgttc	gagcagggac	tegeggtgat	tgtcgátgga	15840
ttggcgaaaa	ggaggctcgt	tgtcaggaac	gttgaaggac	cgagaaaggg	tgacgattga	15900
tcaggaccgc	tgccggagcg	caacccactc	actacagcag	agccatgtag	acaacatccc	15960
ctcccccttt	ccaccgcgtc	agacgcccgt	agcagcccgc	tacgggcttt	ttcatgccct	16020
gccctagcgt	ccaagcctca	cggccgcgct	cggcctctct	ggcggccttc	tggcgctctt	16080
ccgcttcctc	gctcactgac	tegetgeget	cggtcgttcg	gctgcggcga	gcggtatcag	16140
ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	. ggaaagaaca	16200
tgtgagcaaa	ı aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgtttt	16260
tccataggct	: cegececcet	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	16320
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaagctco	: ctcgtgcgct	16380

ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tegggaageg	16440
tggcgctttt	ccgctgcata	accctgcttc	ggggtcatta	tagcgatttt	ttcggtatat	16500
ccatcctttt	tegcaegata	tacaggattt	tgccaaaggg	ttcgtgtaga	ctttccttgg	16560
tgtatccaac	ggcgtcagcc	gggcaggata	ggtgaagtag	gcccacccgc	gagcgggtgt	16620
tecttettea	ctgtccctta	ttcgcacctg	gcggtgctca	acgggaatcc	tgctctgcga	16680
ggctggccgg	ctaccgccgg	cgtaacagat	gagggcaagc	ggatggctga	tgaaaccaag	16740
ccaaccagga	agggcagccc	acctatcaag	gtgtactgcc	ttccagacga	acgaagagcg	16800
attgaggaaa	aggcggcggc	ggccggcatg	agcctgtcgg	cctacctgct	ggccgtcggc	16860
cagggctaca	aaatcacggg	cgtcgtggac	tatgagcacg	tccgcgagct	ggcccgcatc	16920
aatggcgacc	tgggccgcct	gggcggcctg	ctgaaactct	ggctcaccga	cgacccgcgc	16980
acggcgcggt	teggtgatge	cacgatecte	gccctgctgg	cgaagatcga	agagaagcag	17040
gacġagcttg	gcaaggtcat	gatgggcgtg	gtccgcccga	gggcagagcc	atgacttttt	17100
tagccgctaa	aacggccggg	gggtgcgcgt	gattgccaag	cacgtcccca	tgcgctccat	17160
caagaagagc	gacttcgcgg	agctggtgaa	gtacatcacc	gacgagcaag	gcaagaccga	17220
gegeetttge	gacgctca					17238

<210> 40

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<221> misc_feature <222> (3471)..(3471) <223> n is a, c, g, or t <220> <221> misc_feature <222> (3679)..(3679) <223> n is a, c, g, or t <220> <221> misc_feature <222> (3770)..(3770) <223> n is a, c, g, or t <400> 40 gatetttega caetgaaata egtegageet geteegettg gaageggega ggageetegt 60 cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120 cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180 cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240 tegggaagee agegaetggt gaeeteeage gteeetttaa ggetgeeaae agetttetea 300 gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360 gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420 atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480 tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540 aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga 600 gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660 ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 720 tccgatagct ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg 780 tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840

WO 2004/063359 PCT/EP2004/000099 140/357

						000
caccacaaaa	gtcagacggc	gtaaccaaaa	gtcacacaac	acaagctgta	aggatttcgg	900
cacggctacg	gaagacggag	aagccacctt	cagtggactc	gagtaccatt	taattctatt	960
tgtgtttgat	cgagacctaa	tacagcccct	acaacgacca	tcaaagtcgt	atagctacca	1020
gtgaggaagt	ggactcaaat	cgacttcagc	aacatctcct	ggataaactt	taagcctaaa	1080
ctatacagaa	taagataggt	ggagagctta	taccgagctc	ccaaatctgt	ccagatcatg.	1140
gttgaccggt	gcctggatct	tcctatagaa	tcatccttat	tcgttgacct	agctgattct	1200
ggagtgaccc	agagggtcat	gacttgagcc	taaaatccgc	cgcctccacc	atttgtagaa	1260
aaatgtgacg	aactcgtgag	ctctgtacag	tgaccggtga	ctctttctgg	catgcggaga	1320
gacggacgga	cgcagagaga	agggctgagt	aataagccac	tggccagaca	gctctggcgg	1380
ctctgaggtg	cagtggatga	ttattaatcc	gggaccggcc	geceeteege	cccgaagtgg	1440
aaaggctggt	gtgcccctcg	ttgaccaaga	atctattgca	tcatcggaga	atatggagct	1500
tcatcgaatc	accggcagta	agcgaaggag	aatgtgaagc	caggggtgta	tageegtegg	1560
cgaaatagca	tgccattaac	ctaggtacag	aagtccaatt	gcttccgatc	tggtaaaaga	1620
ttcacgagat	agtaccttct	ccgaagtagg	tagagcgagt	acceggegeg	taagctccct	1680
aattggccca	tccggcatct	gtagggcgtc	caaatatcgt	gcctctcctg	ctttgcccgg	1740
tgtatgaaac	cggaaaggcc	gctcaggagc	tggccagcgg	cgcagaccgg	gaacacaagc	1800
tggcagtcga	cccatccggt	gctctgcact	cgacctgctg	aggtccctca	gtccctggta	1860
ggcagctttg	ccccgtctgt	ccgcccggtg	tgtcggcggg	gttgacaagg	tegttgegte	1920
agtccaacat	ttgttgccat	attttcctgd	tctccccacc	agctgctctt	ttettttete	1980
tttcttttcc	: catcttcagt	atattcatct	tcccatccaa	gaacctttat	ttcccctaag	2040
taagtacttt	gctacatcca	tactccatco	: ttcccatccc	ttattccttt	gaacctttca	2100

WO 2004/063359 PCT/EP2004/000099 . 141/357

gttcgagctt	tcccacttca	tcgcagcttg	actaacagct	accccgcttg	agcagacatc	2160
accatgcctg	aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	gttcgacagc	2220
gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	cttcgatgta	2280
ggagggcgtg	gatatgtcct	gcgggtaaat	agctgcgccg	atggtttcta	caaagatcgt	2340
tatgtttatc	ggcactttgc	ateggeegeg	ctcccgattc	cggaagtgct	tgacattggg	2400
gaattcagcg	agagcctgac	ctattgcatc	tcccgccgtg	cacagggtgt	cacgttgcaa	2460
gacctgcctg	aaaccgaact	gcccgctgtt	ctgcagccgg	tegeggagge	catggatgcg	2520
ategetgegg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaatc	2580
ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	tgtgtatcac	2640
tggcaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	cgatgagctg	2700
atgctttggg	ccgaggactg	ccccgaagtc	eggeaceteg	tgcacgcgga	tttcggctcc	2760
aacaatgtcc	tgacggacaa	tggccgcata	acagcggtca	ttgactggag	cgaggcgatg	2820
ttcggggatt	cccaatacga	ggtcgccaac	atettettet	ggaggccgtg	gttggcttgt	2880
atggagcagc	agacgcgcta	cttcgagcgg	aggcatccgg	agcttgcagg	atcgccgcgg	2940
ctccgggcgt	atatgctccg	cattggtctt	gaccaactct	atcagagctt	ggttgacggc	3000
aatttcgatg	atgcagcttg	ggcgcagggt	cgatgcgacg	caatcgtccg	atccggagcc	3060
gggactgtcg	ggcgtacaca	aatcgcccgc	agaagcgcgg	ccgtctggac	cgatggctgt	3120
gtagaagtac	tcgccgatag	tggaaaccga	cgccccagca	ctcgtccgag	ggcaaaggaa	3180
tagagtagat	gccgaccgcg	ggatcgatcc	acttaacgtt	actgaaatca	tcaaacagct	3240
tgacgaatct	ggatataaga	tcgttggtgt	cgatgtcagc	tccggagttg	agacaaatgg	3300
tgttcaggat	: ctcgataaga	tacgttcatt	tgtccaagca	gcaaagagtg	cettetagtg	3360

WO 2004/063359 PCT/EP2004/000099 143/357

gctgcagtca	atcagcgtca	aggcccgccg	cgttgaacta	gcccgcgaca	tcacgcggcc	4680
caaagtctgc	ctgcatgctc	ageggtgete	gttagttcgg	ctgcgagtgg	cagcaccaca	4740
gacagaggag	gegetgggaa	ccgtgcaggc	tgccggcgcg	ggcgatgagc	acagegeega	4800
tgtagcactc	cagcagcttg	accgggctat	cgcagagcgt	cgtgcccggc	gcaaacggga	4860
gcagctgtca	taccaggctg	ccgccattgc	agcatcaatt	ggcgtgtcag	gcattgccat	4920
cttcgccacc	tacctgagat	ttgccatgca	catgaccgtg	ggcggcgcag	tgccatgggg	4980
tgaagtggct	ggcactctcc	tcttggtggt	tggtggcgcg	ctcggcatgg	agatgtatgc	5040
ccgctatgca	cacaaagcca	tctggcatga	gtcgcctctg	ggctggctgc	tgcacaagag	5100
ccaccacaca	cctcgcactg	gaccctttga	agccaacgac	ttgtttgcaa	tcatcaatgg	5160
actgcccgcc	atgctcctgt	gtacctttgg	cttctggctg	cccaacgtcc	tgggggcggc	5220
ctgctttgga	gcggggctgg	gcatcacgct	atacggcatg	gcatatatgt	ttgtacacga	5280
tggcctggtg	cacaggcgct	ttcccaccgg	gcccatcgct	ggcctgccct	acatgaagcg	5340
cctgacagtg	gcccaccagc	tacaccacag	cggcaagtac	ggtggcgcgc	cctggggtat	5400
gttcttgggt	ccacaggagc	tgcagcacat	tccaggtgcg	gcggaggagg	tggagcgact	5460 ·
ggtcctggaa	ctggactggt	ccaagcggta	gattgtgact	gatagcgaga	ctctgggtcg	5520
atgttatctg	cctcaacaat	ggcttagaaa	agaagaaaca	gaacaaatac	agcaaggcaa	5580
cgcccgtagc	ctaggtgatc	aaagactgtt	gggcttgtct	ctgaagcttg	taggaaaggc	5640
agacgctatc	atggtgagag	ctaagaaggg	cattgacaag	ttgccggcaa	actgtcaagg	5700
cggtgtacga	gctgcttgcc	aagtatatgc	tgcaattgga	tctgtactca	agcagcagaa	5760
gacaacatat	cctacaagag	ctcatctaaa	aggaagcgaa	cgtgccaaga	ttgctctgtt	5820
gagtgtatac	aacctctatc	aatctgaaga	caagcctgtg	gctctccgtc	aagctagaaa	5880

WO 2004/063359 PCT/EP2004/000099 144/357

gattaagagt ttttttgttg attagtgaat ttttgtttta tttatgtctg atagttcaat	5940
aaagagacaa cacatacaat ataaaatcat tgtctttaaa tgttaattta gtagagtgta	6000
aageetgeat tttttttgta egeataaaca atgaatteae eeegettetg gtttttaaat	6060
aattatgtca aactagggaa aattetttt tttetetteg ttetttttt ggettgttgt	6120
ggagtcacag gcttgtcttc agattgatag aggttgtata cactcaacag agcaatcttg	6180
gcacgttcgc ttccttttag atgagctctt gtaggatatg ttgtcttctg ctgcttgagt	6240
acagatecaa ttgcageata tacttggcaa geagetegta eacegeettg acagtttgee	6300
ggcaacttgt caatgccctt cttagctctc accatgatag cgtctgcctt tcctacaagc	6360
ttcagagaca agcccaacag tctttgatca cctaggctac gggcgttgcc ttgctgtatt	6420
tgttctgttt cttcttttct aagccattgt tgaggcagat aacatcgacc caacatcctc	6480
gagecatact acageataaa aggatacgtt ttetttaaca gaaatttace ettttgttat	6540
cagcacatac aaaaaaaaag aaatttaaga tgagtaggac ttccattctc tcaaaaattt	6600
tattcaatcc ataaatgaat tatttttgga caaaaaagaa agattatgcc tgattttctc	6660
tattttttt tttttacaa ctccaccaat actttctagc ccagcttggc gtaatcatgg	6720
tcatagctgt ttcctgtgtg aaattgttat ccgctcacaa ttccacacaa catacgagcc	6780
ggaagcataa agtgtaaagc ctggggtgcc taatgagtga gctaactcac attaattgcg	6840
ttgcgctcac tgcccgcttt ccagtcggga aacctgtcgt gccagctgca ttaatgaatc	6900
ggccaacgcg cggggagagg cggtttgcgt attgggccaa agacaaaagg gcgacattca	6960
accgattgag ggagggaagg taaatattga cggaaattat tcattaaagg tgaattatca	7020
ccgtcaccga cttgagccat ttgggaatta gagccagcaa aatcaccagt agcaccatta	7080
ccattagcaa ggccggaaac gtcaccaatg aaaccatcga tagcagcacc gtaatcagta	7140

WO 2004/063359 PCT/EP2004/000099 145/357

	caagtttgcc	tttagggtaa	as atateaca	cattttcatc	ggcattttcg	7200
gcgacagaac	Caageeegee	cctagegeea	gactgtageg	CgCCCCCCC	33-466666	7200
gtcatagccc	ccttattagc	gtttgccatc	ttttcataat	caaaatcacc	ggaaccagag	7260
ccaccaccgg	aaccgcctcc	ctcagagccg	ccaccctcag	aaccgccacc	ctcagagcca	7320
ccaccctcag	agccgccacc	agaaccacca	ccagagccgc	cgccagcatt	gacaggaggc	7380
ccgatctagt	aacatagatg	acaccgcgcg	cgataattta	tcctagtttg	cgcgctatat	7440
tttgttttct	atcgcgtatt	aaatgtataa	ttgcgggact	ctaatcațaa	aaacccatct	7500
cataaataac	gtcatgcatt	acatgttaat	tattacatgc	ttaacgtaat	tcaacagaaa	7560
ttatatgata	atcatcgcaa	gaccggcaac	aggattcaat	cttaagaaac	tttattgcca	7620
aatgtttgaa	cgatcgggga	tcatccgggt	ctgtggcggg	aactccacga	aaatatccga	7680
acgcagcaag	atatcgcggt	gcatctcggt	cttgcctggg	cagtcgccgc	cgacgccgtt	7740
gatgtggacg	cegggcccga	tcatattgtc	gctcaggatc	gtggcgttgt	gcttgtcggc	7800
cgttgctgtc	gtaatgatat	cggcaccttc	gaccgcctgt	tccgcagaga	tcccgtgggc	7860
gaagaactcc	agcatgagat	ccccgcgctg	gaggatcatc	cageeggegt	cccggaaaac	7920
gattccgaag	cccaaccttt	catagaaggc	ggcggtggaa	tcgaaatctc	gtgatggcag	7980
gttgggcgtc	gcttggtcgg	tcatttcgaa	ccccagagtc	ccgctcagaa	gaactcgtca	8040
agaaggcgat	agaaggcgat	gcgctgcgaa	tcgggagcgg	cgataccgta	aagcacgagg	8100
aagcggtcag	cccattcgcc	gccaagctct	tcagcaatat	cacgggtagc	caacgctatg	8160
tcctgatagc	ggtccgccac	acccagccgg	ccacagtcga	tgaatccaga	aaagcggcca	8220
ttttccacca	tgatattcgg	caagcaggca	tegecatggg	tcacgacgag	atcatcgccg	8280
tegggcatge	gcgccttgag	cctggcgaac	agttcggctg	gcgcgagccc	ctgätgctct	8340
tcgtccagat	catcctgatc	gacaagaccg	gcttccatcc	gagtacgtgc	tcgctcgatg	8400

WO 2004/063359 PCT/EP2004/000099 146/357

cgatgtttcg	cttggtggtc	gaatgggcag	gtagccggat	caagcgtatg	cagccgccgc	8460
attgcatcag	ccatgatÿga	tactttctcg	gcaggagcaa	ggtgagatga	caggagatcc	8520
tgccccggca	cttcgcccaa	tagcagccag	tecetteceg	cttcagtgac	aacgtcgagc	8580
acagctgcgc	aaggaacgcc	cgtcgtggcc	agccacgata	gccgcgctgc	ctcgtcctgc	8640
agttcattca	gggcaccgga	caggtcggtc	ttgacaaaaa	gaaccgggcg	cccctgcgct	8700
gacagccgga	acacggcggc	atcagagcag	ccgattgtct	gttgtgccca	gtcatagccg	8760
aatagcctct	ccacccaagc	ggccggagaa	cctgcgtgca	atccatcttg	ttcaatcatg	8820
cgaaacgatc	cagatccggt	gcagattatt	tggattgaga	gtgaatatga	gactctaatt	8880
ggataccgag	gggaatttat	ggaacgtcag	tggagcattt	ttgacaagaa	atatttgcta	8940
gctgatagtg	accttaggcg	acttttgaac	gcgcaataat	ggtttctgac	gtatgtgctt	9000
agctcattaa	actccagaaa	cccgcggctg	agtggctcct	tcaacgttgc	ggttctgtca	9060
gttccaaacg	taaaacggct	tgtcccgcgt	categgeggg	ggtcataacg	tgactccctt	9120 '
aattctccgc	tcatgatcag	attgtcgttt	cccgccttca	gtttaaacta	tcagtgtttg	9180
acaggatata	ttggcgggta	aacctaagag	aaaagagcgt	ttattagaat	aatcggatat	9240
ttaaaagggc	gtgaaaaggt	ttatccgttc	gtccatttgt	atgtgcatgc	caaccacagg	9300
gttccccaga	tctggcgccg	gccagcgaga	cgagcaagat	tggccgccgc	ccgaaacgat	9360
ccgacagcgc	geceageaca	ggtgcgcagg	caaattgcac	caacgcatac	agcgccagca	9420
gaatgccata	ı gtgggcggtg	acgtcgttcg	agtgaaccag	atcgcgcagg	aggcccggca	9480
gcaccggcat	: aatcaggccg	atgccgacag	cgtcgagcgc	gacagtgctc	agaattacga	9540
tcaggggtat	: gttgggtttc	acgtctggcc	: tccggaccag	cctccgctgg	r tccgattgaa	9600
cgcgcggatt	ctttatcact	gataagttgg	g tggacatatt	atgtttatca	gtgataaagt	9660

WO 2004/063359 PCT/EP2004/000099 147/357

gtcaagcatg	acaaagttgc	agccgaatac	agtgatccgt	gccgccctgg	acctgttgaa	9720
cgaggtcggc	gtagacggtc	tgacgacacg	caaactggcg	gaacggttgg	gggttcagca	9780
gccggcgctt	tactggcact	tcaggaacaa	gcgggcgctg	ctcgacgcac	tggccgaagc	9840
catgctggcg	gagaatcata	cgcattcggt	gccgagagcc	gacgacgact	ggcgctcatt	9900
tctgatcggg	aatgcccgca	gcttcaggca	ggegetgete	gcctaccgcg	atggcgcgcg	9960
catccatgcc	ggcacgcgac	cgggcgcacc	gcagatggaa	acggccgacg	cgcagcttcg	10020
cttcctctgc	gaggcgggtt	tttcggccgg	ggacgccgtc	aatgcgctga	tgacaatcag	10080
ctacttcact	gttggggccg	tgcttgagga	gcaggccggc	gacagcgatg	ccggcgagcg	10140
cggcggcacc	gttgaacagg	ctccgctctc	gccgctgttg	cgggccgcga	tagacgcctt	10200
cgacgaagcc	ggtccggacg	cagcgttcga	gcagggactc	gcggtgattg	tcgatggatt	10260
ggcgaaaagg	aggctcgttg	tcaggaacgt	tgaaggaccg	agaaagggtg	acgattgatc	10320
aggaccgctg	ceggagegea	acccactcac	tacagcagag	ccatgtagac	aacatcccct	10380
cccctttcc	accgcgtcag	acgcccgtag	cagcccgcta	cgggcttttt	catgccctgc	10440
cctagcgtcc	aagcctcacg	geegegeteg	gcctctctgg	cggccttctg	gegetettee	10500
gcttcctcgc	tcactgactc	getgegeteg	gtcgttcggc	tgcggcgagc	ggtatcagct	10560
cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	10620
tgagcaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	10680
cataggctcc	gcccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	10740
aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	cgtgcgctct	10800
cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	10860
gcgcttttcc	gctgcataac	cctgcttcgg	ggtcattata	gcgattttt	cggtatatcc	10920

atcettttte geacgatata caggattttg ceaaagggtt egtgtagaet tteettggtg 1098	0
tatccaacgg cgtcagccgg gcaggatagg tgaagtaggc ccacccgcga gcgggtgttc 1104	.0
cttcttcact gtcccttatt cgcacctggc ggtgctcaac gggaatcctg ctctgcgagg 1110	0
ctggccggct accgccggcg taacagatga gggcaagcgg atggctgatg aaaccaagcc 1116	0
aaccaggaag ggcagcccac ctatcaaggt gtactgcctt ccagacgaac gaagagcgat 1122	0:0
tgaggaaaag gcggcggcgg ccggcatgag cctgtcggcc tacctgctgg ccgtcggcca 1128	30
gggctacaaa atcacgggcg tcgtggacta tgagcacgtc cgcgagctgg cccgcatcaa 1134	10
tggcgacctg ggccgcctgg gcggcctgct gaaactctgg ctcaccgacg acccgcgcac 1140	00
ggcgcggttc ggtgatgcca cgatcctcgc cctgctggcg aagatcgaag agaagcagga 1146	50
cgagcttggc aaggtcatga tgggcgtggt ccgcccgagg gcagagccat gactttttta 1152	20
gccgctaaaa cggccggggg gtgcgcgtga ttgccaagca cgtccccatg cgctccatca 115	80
agaagagcga ettegeggag etggtgaagt acateacega egageaagge aagaeegage 116	40
geetttgega egeteacegg getggttgee etegeegetg ggetggegge egtetatgge 117	00
cetgcaaacg egecagaaac geegtegaag eegtgtgega gacaeegegg eegeeggegt 117	60
tgtggatacc tcgcggaaaa cttggccctc actgacagat gaggggcgga cgttgacact 118	20
tgaggggccg actcacccgg cgcggcgttg acagatgagg ggcaggctcg atttcggccg 118	80
gcgacgtgga gctggccagc ctcgcaaatc ggcgaaaacg cctgatttta cgcgagtttc 119	40
ccacagatga tgtggacaag cctggggata agtgccctgc ggtattgaca cttgaggggc 120	00
gcgactactg acagatgagg ggcgcgatcc ttgacacttg aggggcagag tgctgacaga 120	60
tgaggggcgc acctattgac atttgagggg ctgtccacag gcagaaaatc cagcatttgc 121	.20
aagggtttcc gcccgttttt cggccaccgc taacctgtct tttaacctgc ttttaaacca 121	.80

WO 2004/063359 PCT/EP2004/000099 149/357

atatttataa	accttgtttt	taaccagggc	tgcgccctgt	gcgcgtgacc	gcgcacgccg	12240
aaggggggtg	ccccccttc	togaaccctc	ccggcccgct	aacgcgggcc	tcccatcccc	12300
ccaggggctg	cgcccctcgg	ccgcgaacgg	cctcacccca	aaaatggcag	cgctggcagt	12360
ccttgccatt	gccgggatcg	gggcagtaac	gggatgggcg	atcagcccga	gcgcgacgcc	12420
cggaagcatt	gacgtgccgc	aggtgctggc	atcgacattc	agcgaccagg	tgccgggcag	12480
tgagggcggc	ggcctgggtg	geggeetgee	cttcacttcg	gccgtcgggg	cattcacgga	12540
cttcatggcg	gggccggcaa	tttttacctt	gggcattctt	ggcatagtgg	tcgcgggtgc	12600
cgtgctcgtg	ttcgggggtg	cgataaaccc	agcgaaccat	ttgaggtgat	aggtaagatt	12660
ataccgaggt	atgaaaacga	gaattggacc	tttacagaat	tactctatga	agcgccatat	12720
ttaaaaagct	accaagacga	agaggatgaa	gaggatgagg	aggcagattg	ccttgaatat	12780
attgacaata	ctgataagat	aatatatctt	ttatatagaa	gatatcgccg	tatgtaagga	12840
tttcaggggg	caaggcatag	gcagcgcgct	tatcaatata	tctatagaat	gggcaaagca	12900
taaaaacttg	catggactaa	tgcttgaaac	ccaggacaat	aaccttatag	cttgtaaatt	12960
ctatcataat	tgggtaatga	ctccaactta	ttgatagtgt	tttatgttca	gataatgccc	13020
gatgactttg	tcatgcagct	ccaccgattt	tgagaacgac	agcgacttcc	gtcccagccg	13080
tgccaggtgc	tgcctcagat	tcaggttatg	ccgctcaatt	cgctgcgtat	atcgcttgct	13140
gattacgtgc	agctttccct	tcaggcggga	ttcatacagc	ggccagccat	ccgtcatcca	13200
tatcaccacg	tcaaagggtg	acagcaggct	cataagacgc	cccagcgtcg	ccatagtgcg	13260
ttcaccgaat	acgtgcgcaa	caaccgtctt	ccggagactg	tcatacgcgt	aaaacagcca	13320
gegetggege	gatttagccc	cgacatagcc	ccactgttcg	tccatttccg	cgcagacgat	13380
gacgtcactg	cccggctgta	. tgcgcgaggt	taccgactgo	ggcctgagtt	tttaagtga	13440

WO 2004/063359 PCT/EP2004/000099 150/357

cgtaaaatcg	tgttgaggcc	aacgcccata	atgcgggctg	ttgcccggca	tccaacgcca	13500
ttcatggcca	tatcaatgat	tttctggtgc	gtaccgggtt	gagaagcggt	gtaagtgaac	13560
tgcagttgcc	atgttttacg	gcagtgagag	cagagatagc	gctgatgtcc	ggcggtgctt	13620
ttgccgttac	gcaccacccc	gtcagtagct	gaacaggagg	gacagctgat	agacacagaa	13680 .
gccactggag	cacctcaaaa	acaccatcat	acactaaatc	agtaagttgg	cagcatcacc	13740
cataattgtg	gtttcaaaat	cggctccgtc	gatactatgt	tatacgccaa	ctttgaaaac	13800
aactttgaaa	aagctgtttt	ctggtattta	aggttttaga	atgcaaggaa	cagtgaattg	13860
gagttcgtct	tgttataatt	agcttcttgg	ggtatcttta	aatactgtag	aaaagaggaa	13920
ggaaataata	aatggctaaa	atgagaatat	caccggaatt	gaaaaaactg	atcgaaaaat	13980
accgctgcgt	aaaagatacg	gaaggaatgt	ctcctgctaa	ggtatataag	ctggtgggag	14040
aaaatgaaaa	cctatattta	aaaatgacgg	acagccggta	taaagggacc	acctatgatg	14100
tggaacggga	aaaggacatg	atgctatggc	tggaaggaaa	gctgcctgtt	ccaaaggtcc	14160
tgcactttga	acggcatgat	ggctggagca	atctgctcat	gagtgaggcc	gatggcgtcc	14220
tttgctcgga	agagtatgaa	gatgaacaaa	gccctgaaaa	gattatcgag	ctgtatgcgg	14280
agtgcatcag	gctctttcac	tccatcgaca	tatcggattg	tccctatacg	aatagcttag	14340
acageegett	agccgaattg	gattacttac	tgaataacga	tctggccgat	gtggattgcg	14400
aaaactggga	agaagacact	ccatttaaag	atccgcgcga	gctgtatgat	tttttaaaga	14460
cggaaaagcc	cgaagaggaa	cttgtcttt	cccacggcga	cctgggagac	agcaacatct	14520
ttgtgaaaga	tggcaaagta	agtggcttta	ttgatcttgg	gagaagcggc	agggcggaca	14580
agtggtatga	cattgccttc	tgcgtccggt	cgatcaggga	ggatateggg	gaagaacagt	14640
atgtcgagct	attttttgac	ttactgggga	tcaagcctga	ttgggagaaa	ataaaatatt	14700

WO 2004/063359 PCT/EP2004/000099 151/357

atattttact	ggatgaattg	ttttagtacc	tagatgtggc	gcaacgatgc	cggcgacaag	1:4760
caggagcgca	ccgacttctt	ccgcatcaag	tgttttggct	ctcaggccga	ggcccacggc	14820
aagtatttgg	gcaaggggtc	gctggtattc	gtgcagggca	agattcggaa	taccaagtac	14880
gagaaggacg	gccagacggt	ctacgggacc	gacttcattg	ccgataaggt	ggattatctg	14940
gacaccaagg	caccaggcgg	gtcaaatcag	gaataagggc	acattgcccc	ggcgtgagtc	15000
ggggcaatcc	cgcaaggagg	gtgaatgaat	cggacgtttg	accggaaggc	ata'caggcaa	15060
gaactgatcg	acgcggggtt	ttccgccgag	gatgccgaaa	ccatcgcaag	ccgcaccgtc _.	15120 .
atgcgtgcgc	cccgcgaaac	cttccagtcc	gtcggctcga	tggtccagca	agctacggcc	15180
aagatcgagc	gcgacagcgt	gcaactggct	cccctgccc	tgcccgcgcc	atcggccgcc	15240
gtggagcgtt	cgcgtcgtct	cgaacaggag	gcggcaggtt	tggcgaagtc	gatgaccatc	15300
gacacgcgag	gaactatgac	gaccaagaag	cgaaaaaccg	ccggcgagga	cctggcaaaa	15360
caggtcagcg	aggccaagca	ggccgcgttg	ctgaaacaca	cgaagcagca	gatcaaggaa	15420
atgcagcttt	ccttgttcga	tattgcgccg	tggccggaca	cgatgcgagc	gatgccaaac	15480
gacacggccc	gatatgacat	gttcaccacg	cgcaacaaga	aaatcccgcg	cgaggcgctg	15540
caaaacaagg	tcattttcca	cgtcaacaag	gacgtgaaga	tcacctacac	cggcgtcgag	15600
ctgcgggccg	acgatgacga	actggtgtgg	cagcaggtgt	tggagtacgc	gaagcgcacc	15660
cctatcggcg	agccgatcac	cttcacgttc	tacgagcttt	gccaggacct	gggctggtcg	15720
atcaatggco	: ggtattacac	gaaggccgag	gaatgcctgt	. egegeetaca	ggcgacggcg	15780
atgggcttca	ı cgtccgaccg	cgttgggcac	ctggaatcgg	tgtegetget	gcaccgcttc	15840
cgcgtcctgg	g accgtggcaa	gaàaacgtcc	e cgttgccagg	, teetgatega	cgaggaaatc	15900
gtcgtgctgt	: ttgctggcga	ccactacacg	aaattcatat	gggagaagta	ccgcaagctg	15960

WO 2004/063359 PCT/EP2004/000099 152/357

tegeegaegg ceegaeggat gttegaetat tteagetege acegggagee gtaeeegete 16020 aagetggaaa cetteegeet catgtgegga teggatteea eeegegtgaa gaagtggege 16080 gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 16140 gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 16200 ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 16260 tgcttcgctc agtatcgctc gggacgcacg gcgcgctcta cgaactgccg ataaacagag 16320 gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 16380 aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 16440 tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 16500 tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 16560 acggccccaa ggacgctcac aaggcgcatc tgtccggcgt tttcgtggag cccgaacagc 16620 gaggccgagg ggtcgccggt atgctgctgc gggcgttgcc ggcgggttta ttgctcgtga 16680 tgatcgtccg acagattcca acgggaatct ggtggatgcg catcttcatc ctcggcgcac 16740 ttaatatttc gctattctgg agcttgttgt ttatttcggt ctaccgcctg ccgggcgggg 16800 tegeggegae ggtaggeget gtgcageege tgatggtegt gttcatetet geegetetge 16860 taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 16920 cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 16980 cggggggggt ttccatggcg ttcggaaccg tgctgacccg caagtggcaa cctcccgtgc 17040 ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 17100 ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 17160 ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 17220

WO 2004/063359 PCT/EP2004/000099 153/357

tegaacetae	agttgtttcc	ttactgggct	ttctcagccc	cagatctggg	gtcgatcagc	17280
cggggatgca	tcaggccgac	agtcggaact	tegggteecc	gacctgtacc	attcggtgag	17340
caatggatag	gggagttgat	atcgtcaacg	ttcacttcta	aagaaatagc	gccactcagc	17400
ttcctcagcg	gctttatcca	gcgatttcct	attatgtcgg	catagttctc	aagatcgaca	17460
gcctgtcacg	gttaagcgag	aaatgaataa	gaaggctgat	aattcggatc	tctgcgaggg	17520
agatgatatt	tgatcacagg	cagcaacgct	ctgtcatcgt	tacaatcaac	atgctaccct	17580
ccgcgagatc	atccgtgttt	caaacccggc	agcttagttg	ccgttcttcc	gaatagcatc	17640
ggtaacatga	gcaaagtctg	ccgccttaca	acggctctcc	cgctgacgcc	gtcccggact	17700
gatgggctgc	ctgtatcgag	tggtgatttt	gtgccgagct	gccggtcggg	gagctgttgg	17760
ctggctggtg	gcaggatata	ttgtggtgta	aacaaattga	cgcttagaca	acttaataac	17820
acattgcgga	cgtttttaat	gtactggggt	ggtttttctt	ttcaccagtg	agacgggcaa	17880
cagctgattg	cccttcaccg	cctggccctg	agagagttgc	agcaagcggt	ccacgctggt	17940
ttgccccagc	aggcgaaaat	cctgtttgat	ggtggttccg	aaatcggcaa	aatcccttat	18000
aaatcaaaag	aatagcccga	gatagggttg	agtgttgttc	cagtttggaa	caagagtcca	18060
ctattaaaga	acgtggactc	caacgtcaaa	gggcgaaaaa	ccgtctatca	gggcgatggc	18120
ccactacgtg	aaccatcacc	caaatcaagt	tttttggggt	cgaggtgccg	taaagcacta	18180
aatcggaacc	ctaaagggag	ccccgattt	agagcttgac	ggggaaagco	ggegaaegtg	18240
gcgagaaagg	aagggaagaa	agcgaaagga	gcgggcgcca	ttcaggctgc	gcaactgttg:	18300
ggaagggcga	teggtgeggg	cctcttcgct	attacgccag	ctggcgaaag	ggggatgtgc	18360
tgcaaggcga	ttaagttggg	taacgccagg	gttttcccag	tcacgacgtt	: gtaaaacgac	18420
ggccagtgaa	ttcgagctcg	gtacccggg				18449

WO 2004/063359 PCT/EP2004/000099

<210> 41 <211> 18449 <212> DNA <213> Artificial <220> <223> Plasmid <220> <221> misc_feature <222> (3471)..(3471) <223> n is a, c, g, or t <220> <221> misc_feature <222> (3679)..(3679) <223> n is a, c, g, or t <220> <221> misc_feature <222> (3770)..(3770) <223> n is a, c, g, or t <400> 41 gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120 cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180 240 cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct tcgggaagcc agcgactggt gacctccagc gtccctttaa ggctgccaac agctttctca 300 gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360 420 gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca 480

WO 2004/063359 PCT/EP2004/000099 155/357

tttggatttg	ctttccaggc	tgagactcta	gcttggagca	tagagggtcc	tttggctttc	540
aatattctca	agtatctcga	gtttgaactt	attccctgtg	aaccttttat	tcaccaatga	600
gcattggaat	gaacatgaat	ctgaggactg	caatcgccat	gaggttttcg	aaatacatcc	660
ggatgtcgaa	ggcttggggc	acctgcgttg	gttgaattta	gaacgtggca	ctattgatca	720
tccgatagct	ctgcaaaggg	cgttgcacaa	tgcaagtcaa	acgttgctag	cagttccagg	780
tggaatgtta	tgatgagcat	tgtattaaat	caggagatat	agcatgatct	ctagttagct	840
caccacaaaa	gtcagacggc	gtaaccaaaa	gtcacacaac	acaagctgta	aggatttcgg	900
cacggctacg	gaagacggag	aagccacctt	cagtggactc	gagtaccatt	taattctatt	960
tgtgtttgat	cgagacctaa	tacagcccct	acaacgacca	tcaaagtcgt	atagctacca	1020
gtgaggaagt	ggactcaaat	cgacttcagc	aacatctcct	ggataaactt	taagcctaaa	1080
ctatacagaa	taagataggt	ggagagctta	taccgagctc	ccaaatctgt	ccagatcatg	1140
gttgaccggt	gcctggatct	tcctatagaa	tcatccttat	tegttgacet	agctgattct	1200
ggagtgaccc	agagggtcat	gacttgagcc	taaaatccgc	egeeteeace	atttgtagaa	1260
aaatgtgacg	aactcgtgag	ctctgtacag	tgaccggtga	ctctttctgg	catgcggaga	1320
gacggacgga	cgcagagaga	agggctgagt	aataagccac	tggccagaca	gctctggcgg	1380
ctctgaggtg	cagtggatga	ttattaatcc	gggaccggcc	gcccctccgc	cccgaagtgg	1440
aaaggctggt	gtgcccctcg	ttgaccaaga	atctattgca	tcatcggaga	atatggagct	1500
tcatcgaatc	accggcagta	agcgaaggag	aatgtgaagc	caggggtgta	tagccgtcgg	1560
cgaaatagca	. tgccattaac	ctaggtacag	aagtccaatt	gcttccgatc	tggtaaaaga	1620
ttcacgagat	. agtaccttct	ccgaagtagg	tagagcgagt	acceggegeg	taagctccct	1680
aattggccca	teeggeatet	gtagggcgtc	caaatatcgt	geeteteetg	ctttgcccgg	1740

WO 2004/063359 PCT/EP2004/000099 156/357

tgtatgaaac	cggaaaggcc	gctcaggagc	tggccagcgg	cgcagaccgg	gaacacaagc	1800
tggcagtcga	cccatccggt	gctctgcact	cgacctgctg	aggtccctca	gtccctggta	1860
ggcagctttg	ccccgtctgt	ccgcccggtg	tgtcggcggg	gttgacaagg	tcgttgcgtc	1920
agtccaacat	ttgttgccat	attttcctgc	tetececace	agetgetett	ttcttttctc	1980
tttcttttcc	catcttcagt	atattcatct	teccatecaa	gaacctttat	ttcccctaag	2040
taagtacttt	gctacatcca	tactccatcc	ttcccatccc	ttattccttt	gaacctttca	2100
gttcgagctt	tcccacttca	tegeagettg	actaacagct	accccgcttg	agcagacatc	2160
accatgcctg	aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	gttcgacagc	2220
gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	cttcgatgta	2280
ggagggcgtg	gatatgtcct	gcgggtaaat	agctgcgccg	atggtttcta	caaagatcgt	2340
tatgtttatc	ggcactttgc	atcggccgcg	ctcccgattc	cggaagtgct	tgacattggg	2400
gaattcagcg	agagcctgac	ctattgcatc	tecegeegtg	cacagggtgt	cacgttgcaa	2460
gacctgcctg	aaaccgaact	gcccgctgtt	ctgcagccgg	tcgcggaggc	catggatgcg	2520
atcgctgcgg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaatc	2580
ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	tgtgtatcac	2640
tggcaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	cgatgagctg	2700
atgctttggg	ccgaggactg	ccccgaagtc	cggcacctcg	tgcacgcgga	tttcggctcc	2760
aacaatgtcc	tgacggacaa	tggccgcata	acageggtca	ttgactggag	cgaggcgatg	2820
ttcggggatt	cccaatacga	ggtcgccaac	atcttcttct	ggaggccgtg	gttggcttgt	2880
atggagcagc	agacgcgcta	cttcgagcgg	aggcateegg	agcttgcagg	ategeegegg	2940
ctccgggcgt	atatgctccg	cattggtctt	gaccaactct	atcagagctt	ggttgacggc	3000

WO 2004/063359 PCT/EP2004/000099 157/357

aatttcgatg	atgcagcttg	ggcgcagggt	cgatgcgacg	caatcgtccg	atccggagcc	3060
gggactgtcg	ggcgtacaca	aatcgcccgc	agaagcgcgg	ccgtctggac	cgatggctgt	3120
gtagaagtac	tcgccgatag	tggaaaccga	cgccccagca	ctcgtccgag	ggcaaaggaa	3180
tagagtagat	gccgaccgcg	ggatcgatcc	acttaacgtt	actgaaatca	tcaaacagct	3240
tgacgaatct	ggatataaga	tcgttggtgt	cgatgtcagc	tccggagttg	agacaaatgg	3300
tgtťcaggat	ctcgataaga	tacgttcatt	tgtccaagca	gcaaagagtg	ccttctagtg	3360
atttaatagc	tccatgtcaa	caagaataaa	acgcgttttc	gggtttacct	cttccagata	3420
cagctcatct	gcaatgcatt	aatgcattga	ctgcaaccta	gtaacgcctt	ncaggctccg	3480
gcgaagagaa	gaatagctta	gcagagctat	tttcattttc	gggagacgag	atcaagcaga	3540
tcaacggtcg	tcaagagacc	tacgagactg	aggaatccgc	tettggetee	acgcgactat	3600
atatttgtct	ctaattgtac	tttgacatgc	tectettett	tactctgata	gcttgactat	3660
gaaaattccg	tcaccagene	ctgggttcgc	aaagataatt	gcatgtttct	tccttgaact	3720
ctcaagccta	ı caggacacac	attcatcgta	ggtataaacc	tcgaaatcan	ttcctactaa	3780
gatggtatac	: aatagtaacc	atgcatggtt	gcctagtgaa	tgctccgtaa	cacccaatac	3840
gccggccgaa	actttttac	aactctccta	tgagtcgttt	acccagaatg	r cacaggtaca	3900
cttgtttaga	a ggtaatcctt	: ctttctagct	agaagteete	gtgtactgtg	, taagcgccca	3960
ctccacatc	t ccactcgacc	: tgcaggcatg	caaagcttga	ı gattaaaata	a gataaggaaa	4020
agaaagtga	a aagaaattcg	g gaagcatggc	acattettet	: ttttataaat	acatgcctga	4080
ctttcttt	t ccatcgatat	t gatatatgca	ı tatgatagat	: atacaagcaa	a tottottcaa	4140
ggagtttga	a attttgtcct	t ccaggagcaa	a aaaaaagttt	: ttttttata	c atgtttgtac	4200
acaagaata	g ttaccaatt	t gctttggtct	tacgtgctgc	c aagtttata	t cgttttcaat	4260

WO 2004/063359 PCT/EP2004/000099 158/357

ttctttgtct	ttacattttc	tttgtccttt	atctttcctc	atttagtctt	tgggagaatt	4320
aggaaaaggg	agcggaaagg	taagaaatgc	ttgcgtattt	tactaattcg	gcaaacatcc	4380
aatttggcaa	acagcagcct	gtgcaacgct	ctcgagatga	cagtatcttt	gattacactc	4440
taaatctcga	tgacccgacc	aaaaagagcg	aacaaagaaa	taatcttgtg	cattcgaata	4500
tgatggaaga	tttttcccc	cttattctaa	atgttgacat	agcgtgtatg	ttatataaac	4560
aaaaagaaat	tgtacaaact	ttctttctt	ctctttttat	tttatctcta	tgctgtcgaa	4620
gctgcagtca	atcagcgtca	aggcccgccg	cgttgaacta	gcccgcgaca	tcacgcggcc	4680
caaagtctgc	ctgcatgctc	agcggtgctc	gttagttcgg	ctgcgagtgg	cagcaccaca	4740
gacagaggag	gcgctgggaa	ccgtgcaggc	tgccggcgcg	ggcgatgagc	acagcgccga	4800
tgtagcactc	cagcagcttg	accgggctat	cgcagagcgt	cgtgcccggc	gcaaacggga	4860
gcagctgtca	taccaggctg	ccgccattgc	agcatcaatt	ggcgtgtcag	gcattgccat	4920
cttcgccacc	tacctgagat	ttgccatgca	catgaccgtg	ggcggcgcag	tgccatgggg	4980
tgaagtggct	ggcactctcc	tcttggtggt	tggtggcgcg	ctcggcatgg	agatgtatgc	5040
ccgctatgca	cacaaagcca	tctggcatga	gtcgcctctg	ggctggctgc	tgcacaagag	5100
ccaccacaca	cctcgcactg	gaccctttga	agccaacgac	ttgtttgcaa	tcatcaatgg	5160
actgcccgcc	atgctcctgt	gtacctttgg	cttctggctg	cccaacgtcc	tgggggcggc	5220
ctgctttgga	gcggggctgg	gcatcacgct	atacggcatg	gcatatatgt	ttgtacacga	5280
tggcctggtg	cacaggeget	ttcccaccgg	gcccatcgct	ggcctgccct	acatgaagcg	5340
cctgacagtg	gcccaccagc	tacaccacag	cggcaagtac	ggtggcgcgc	cctggggtat	5400
gttcttgggt	ccacaggagc	tgcagcacat	tccaggtgcg	gcggaggagg	tggagcgact	5460
ggtcctggaa	ctggactggt	ccaagcgggc	gattgtgact	gatagcgaga	ctctgggtcg	5520

WO 2004/063359 PCT/EP2004/000099 159/357

atgttatctg	cctcaacaat	ggcttagaaa	agaagaaaca	gaacaaatac	agcaaggcaa	5580
cgcccgtagc	ctaggtgatc	aaagactgtt	gggcttgtct	ctgaagcttg	taggaaaggc	5640
agacgctatc	atggtgagag	ctaagaaggg	cattgacaag	ttgccggcaa	actgtcaagg	5700
cggtgtacga	gctgcttgcc	aagtatatgc	tgcaattgga	tctgtactca	agcagcagaa	5760
gacaacatat	cctacaagag	ctcatctaaa	aggaagcgaa	cgtgccaaga	ttgctctgtt	5820
gagtgtatac	aacctctatc	aatctgaaga	caagcctgtg	gctctccgtc	aagctagaaa	5880
gattaagagt	ttttttgttg	attagtgaat	ttttgtttta	tttatgtctg	atagttcaat	5940
aaagagacaa	cacatacaat	ataaaatcat	tgtctttaaa	tgttaattta	gtagagtgta	6000
aagcctgcat	tttttttgta	cgcataaaca	atgaattcac	cccgcttctg	gtttttaaat	6060
aattatgtca	aactagggaa	aattctttt	tttctcttcg	ttctttttt	ggcttgttgt	6120
ggagtcacag	gcttgtcttc	agattgatag	aggttgtata	cactcaacag	agcaatcttg	6180
gcacgttcgc	ttccttttag	atgagctctt	gtaggatatg	ttgtcttctg	ctgcttgagt	6240
acagatccaa	ttgcagcata	tacttggcaa	gcagctcgta	caccgccttg	acagtttgcc	6300
ggcaacttgt	caatgccctt	cttagctctc	accatgatag	cgtctgcctt	tcctacaagc	6360
ttcagagaca	agcccaacag	tctttgatca	cctaggctac	gggcgttgcc	ttgctgtatt	6420
tgttctgttt	cttcttttct	aagccattgt	tgaggcagat	aacatcgacc	caacatcctc	6480
gagccatact	acagcataaa	aggatacgţt	ttctttaaca	gaaatttacc	cttttgttat	6540
cagcacatac	aaaaaaaag	aaatttaaga	tgagtaggac	ttccattctc	tcaaaaattt	6600
tattcaatcc	ataaatgaat	tatttttgga	caaaaaagaa	agattatgcc	tgattttctc	6660
tattttttt	tttttacaa	ctccaccaat	actttctago	ccagettgge	gtaatcatgg	6720
tcatagctgt	ttcctgtgtg	aaattgttat	ccgctcacaa	ttccacacaa	catacgagcc	6780

WO 2004/063359 PCT/EP2004/000099 160/357

ggaagcataa	agtgtaaagc	ctggggtgcc	taatgagtga	gctaáctcac	attaattgcg	6840
ttgcgctcac	tgcccgcttt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	6900
ggccaacgcg	cggggagagg	cggtttgcgt	attgggccaa	agacaaaagg	gcgacattca	6960
accgattgag	ggagggaagg	taaatattga	cggaaattat	tcattaaagg	tgaattatca	7020
ccgtcaccga	cttgagccat	ttgggaatta	gagccagcaa	aatcaccagt	agcaccatta	7080
ccattagcaa	ggccggaaac	gtcaccaatg	aaaccatcga	tagcagcacc	gtaatcagta	7140
gcgacagaat	caagtttgcc	tttagcgtca	gactgtagcg	cgttttcatc	ggcattttcg	7200
gtcatagccc	ccttattagc	gtttgccatc	ttttcataat	caaaatcacc	ggaaccagag	7260
ccaccaccgg	aaccgcctcc	ctcagagccg	ccaccctcag	aaccgccacc	ctcagagcca	7320
ccaccctcag	ageegeeace	agaaccacca	ccagagccgc	cgccagcatt	gacaggaggc	7380
ccgatctagt	aacatagatg	acaccgcgcg	cgataattta	tcctagtttg	cgcgctatat	7440
tttgttttct	atcgcgtatt	aaatgtataa	ttgcgggact	ctaatcataa	aaacccatct	7500
cataaataac	gtcatgcatt	acatgttaat	tattacatgc	ttaacgtaat	tcaacagaaa	7560
ttatatgata	atcatcgcaa	gaccggcaac	aggattcaat	cttaagaaac	tttattgcca	7620
aatgtttgaa	cgatcgggga	tcatccgggt	ctgtggcggg	aactccacga	aaatatccga	7680
acgcagcaag	atatcgcggt	gcatctcggt	cttgcctggg	cagtcgccgc	cgacgccgtt	7740
gatgtggacg	ccgggcccga	tcatattgtc	gctcaggatc	gtggcgttgt	gcttgtcggc	7800
cgttgctgtc	gtaatgatat	cggcaccttc	gaccgcctgt	tccgcagaga	tcccgtgggc	7860
gaagaactco	: agcatgagat	. ccccgcgctg	gaggatcato	cageeggegt	cccggaaaac	7920
gattccgaag	g cccaaccttt	. catagaaggc	ggcggtggaa	tcgaaatctc	gtgatggcag	7980
gttgggcgto	gettggtegg	tcatttcgaa	cccagagto	: ccgctcagaa	gaactcgtca	8040

WO 2004/063359 PCT/EP2004/000099

agaaggcgat agaaggcgat gcgctgcgaa tcgggagcgg cgataccgta aagcacgagg 8100 aagcggtcag cccattcgcc gccaagctct tcagcaatat cacgggtagc caacgctatg 8160 8220 tcctgatage ggtccgccae acccagccgg ccacagtcga tgaatccaga aaagcggcca 8280 ttttccacca tgatattcgg caagcaggca tcgccatggg tcacgacgag atcatcgccg tegggeatge gegeettgag eetggegaae agtteggetg gegegageee etgatgetet 8340 8400 tegtecagat cateetgate gacaagaceg gettecatee gagtacgtge tegetegatg 8460 cgatgtttcg cttggtggtc gaatgggcag gtagccggat caagcgtatg cagccgccgc 8520 attgcatcag ccatgatgga tactttctcg gcaggagcaa ggtgagatga caggagatcc 8580 tgccccggca cttcgcccaa tagcagccag tcccttcccg cttcagtgac aacgtcgagc acagetgege aaggaaegee egtegtggee agecaegata geegegetge etegteetge 8640 8700 agttcattca gggcaccgga caggtcggtc ttgacaaaaa gaaccgggcg cccctgcgct 8760 gacageegga acaeggegge ateagageag eegattgtet gttgtgeeca gteatageeg 8820 aatagcctct ccacccaagc ggccggagaa cctgcgtgca atccatcttg ttcaatcatg 8880 cgaaacgatc cagatccggt gcagattatt tggattgaga gtgaatatga gactctaatt 8940 ggataccgag gggaatttat ggaacgtcag tggagcattt ttgacaagaa atatttgcta 9000 gctgatagtg accttaggcg acttttgaac gcgcaataat ggtttctgac gtatgtgctt 9060 ageteattaa aeteeagaaa eeegeggetg agtggeteet teaaegttge ggttetgtea gttccaaacg taaaacggct tgtcccgcgt catcggcggg ggtcataacg tgactccctt 9120 aatteteege teatgateag attgtegttt ceegeettea gtttaaacta teagtgtttg 9180 9240 acaggatata ttggcgggta aacctaagag aaaagagcgt ttattagaat aatcggatat 9300 ttaaaagggc gtgaaaaggt ttatccgttc gtccatttgt atgtgcatgc caaccacagg

WO 2004/063359 PCT/EP2004/000099 162/357

gttccccaga	tetggegeeg	gccagcgaga	cgagcaagat	tggccgccgc	ccgaaacgat	9360
		ggtgcgcagg				9420
						9480
•		acgtcgttcg				
gcaccggcat	aatcaggccg	atgccgacag	cgtcgagcgc	gacagtgctc	agaattacga	9540
tcaggggtat	gttgggtttc	acgtctggcc	teeggaceag	cctccgctgg	tccgattgaa	9600
cgcgcggatt	ctttatcact	gataagttgg	tggacatatt	atgtttatca	gtgataaagt	9660
gtcaagcatg	acaaagttgc	agccgaatac	agtgatccgt	gccgccctgg	acctgttgaa	9720
cgaggtcggc	gtagacggtc	tgacgacacg	caaactggcg	gaacggttgg	gggttcagca	9780
gccggcgctt	tactggcact	tcaggaacaa	gcgggcgctg	ctcgacgcac	tggccgaagc	9840
catgctggcg	gagaatcata	cgcattcggt	gccgagagcc	gacgacgact	ggcgctcatt	9900
tctgatcggg	aatgcccgca	gcttcaggca	ggcgctgctc	gcctaccgcg	atggcgcgcg	9960
catccatgcc	ggcacgcgac	cgggcgcacc	gcagatggaa	acggccgacg	cgcagcttcg	10020
cttcctctgc	gaggcgggtt	ttteggeegg	ggacgccgtc	aatgcgctga	tgacaatcag	10080
ctacttcact	gttggggccg	tgcttgagga	gcaggccggc	gacagcgatg	ccggcgagcg	10140
cggcggcacc	gttgaacagg	ctecgetete	gccgctgttg	cgggccgcga	tagacgcctt	10200
cgacgaagcc	ggtccggacg	cagcgttcga	gcagggactc	gcggtgattg	tcgatggatt	10260
ggcgaaaagg	aggctcgttg	tcaggaacgt	tgaaggaccg	agaaagggtg	acgattgatc	10320
aggaccgctg	r ccggagcgca	acccactcac	tacagcagag	ccatgtagac	aacatcccct	10380
cccctttcc	: accgcgtcag	acgcccgtag	cagecegeta	egggetttt	. catgccctgc	10440
cctagcgtcc	aagcctcacg	geegegeteg	gcctctctgg	cggccttctg	gegetettee	10500
gattactaga	tcactgactc	: gctgcgctcg	gtcgttcggc	tgcggcgago	ggtatcagct	10560

WO 2004/063359 PCT/EP2004/000099 163/357

(cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	10620
1	tgagcaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	10680
			cgagcatcac				10740
	aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaagctccct	egtgegetet	10800
	cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	10860
,	gcgcttttcc	gctgcataac	cctgcttcgg	ggtcattata	gcgattttt	cggtatatcc	10920
	atccttttc	gcacgațata	caggattttg	ccaaagggtt	cgtgtagact	ttccttggtg	10980
	tatccaacgg	cgtcagccgg	gcaggatagg	tgaagtaggc	ccacccgcga	gcgggtgttc	11040
	cttcttcact	gtcccttatt	cgcacctggc	ggtgctcaac	gggaatcctg	ctctgcgagg	11100
	ctggccggct	accgccggcg	taacagatga	gggcaagcgg	atggctgatg	aaaccaagcc	11160
	aaccaggaag	ggcagcccac	ctatcaaggt	gtactgcctt	ccagacgaac	gaagagcgat	11220
	tgaggaaaag	gċggcggcgg	ccggcatgag	cctgtcggcc	tacctgctgg	ccgtcggcca	11280
	gggctacaaa	atcacgggcg	tcgtggacta	tgagcacgtc	cgcgagctgg	cccgcatcaa	11340
	tggcgacctg	ggeegeetgg	geggeetget	gaaactctgg	ctcaccgacg	acccgcgcac	11400
	ggcgcggttc	ggtgatgcca	cgatectege	cctgctggcg	aagatcgaag	agaagcagga	11460
	cgagcttggc	aaggtcatga	tgggcgtggt	ccgcccgagg	gcagagccat	gacttttta	11520
	gccgctaaaa	cggccgggg	gtgcgcgtga	ttgccaagca	cgtccccatg	cgctccatca	11580
	agaagagcga	ı cttcgcggag	ctggtgaagt	acatcaccga	ı cgagcaaggo	aagaccgagc	11640
	gcctttgcga	ı cgctcaccgg	getggttgee	: ctcgccgctg	ggctggcggc	: cgtctatggc	11700
	cctgcaaacg	g cgccagaaac	geegtegaag	g ccgtgtgcga	ı, gacaccgcgg	cegeeggegt	11760
	tgtggataco	c togoggaaaa	a cttggcccto	c actgacagat	gaggggcgga	a cgttgacact	11820

WO 2004/063359 PCT/EP2004/000099 164/357

tgaggggccg	actcacccgg	cgcggcgttg	acagatgagg	ggcaggctcg	atttcggccg	11880
gcgacgtgga	gctggccagc	ctcgcaaatc	ggcgaaaacg	cctgatttta	cgcgagtttc	11940
ccacagatga	tgtggacaag	cctggggata	agtgccctgc	ggtattgaca	cttgaggggc	12000
gcgactactg	acagatgagg	ggcgcgatcc	ttgacacttg	aggggcagag	tgctgacaga	12060
tgaggggcgc	acctattgac	atttgagggg	ctgtccacag	gcagaaaatc	cagcatttgc	12120
aagggtttcc	gcccgttttt	cggccaccgc	taacctgtct	tttaacctgc	ttttaaacca	12180
atatttataa	accttgtttt	taaccagggc	tgcgccctgt	gegegtgace	gcgcacgccg	12240
aaggggggtg	ccccccttc	tegaaceete	ceggeceget	aacgcgggcc	teccatecee	12300
ccaggggctg	cgcccctcgg	ccgcgaacgg	cctcacccca	aaaatggcag	cgctggcagt	12360
ccttgccatt	gccgggatcg	gggcagtaac	gggatgggcg	atcagcccga	gcgcgacgcc	12420
cggaagcatt	gacgtgccgc	aggtgctggc	atcgacattc	agcgaccagg	tgccgggcag	12480
tgagggcggc	ggcctgggtg	geggeetgee	cttcacttcg	gccgtcgggg	cattcacgga	12540
cttcatggcg	gggccggcaa	tttttacctt	gggcattctt	ggcatagtgg	tegegggtge	12600
cgtgctcgtg	ttcgggggtg	cgataaaccc	agcgaaccat	ttgaggtgat	aggtaagatt	12660
ataccgaggt	. atgaaaacga	gaattggacc	tttacagaat	tactctatga	agcgccatat	12720
ttaaaaagct	accaagacga	agaggatgaa	gaggatgagg	aggcagattg	ccttgaatat	12780
attgacaata	ı ctgataagat	aatatatctt	: ttatatagaa	gatatcgccg	, tatgtaagga	12840
tttcaggggg	g caaggcatag	gcagcgcgct	: tatcaatata	tctatagaat	gggcaaagca	12900
taaaaactto	g catggactaa	ı tgcttgaaac	c ccaggacaat	aaccttatag	g cttgtaaatt	12960
ctatcataat	: tgggtaatga	ı ctccaactta	ı ttgatagtgt	: tttatgttca	a gataatgece	13020
gatgacttt	g tcatgcagct	: ccaccgattt	tgagaacgad	agegaette	gtcccagccg	13080

WO 2004/063359 PCT/EP2004/000099 165/357

tgccaggtgc	tgcctcagat	tcaggttatg	ccgctcaatt	cgctgcgtat	atcgcttgct	13140
gattacgtgc	agctttccct	tcaggcggga	ttcatacagc	ggccagccat	ccgtcatcca	13200
tatcaccacg	tcaaagggtg	acagcaggct	cataagacgc	cccagcgtcg	ccatagtgcg	13260
ttcaccgaat	acgtgcgcaa	caaccgtctt	ccggagactg	tcatacgcgt	aaaacagcca	13320
gegetggege	gatttagccc	cgacatagcc	ccactgttcg	tccatttccg	cgcagacgat	13380
gacgtcactg	cccggctgta	tgcgcgaggt	taccgactgc	ggcctgagtt	ttttaagtga	13440
cgtaaaatcg	tgttgaggcc	aacgcccata	atgegggetg	ttgcccggca	tccaacgcca	13500
ttcatggcca	tatcaatgat	tttctggtgc	gtaccgggtt	gagaagcggt	gtaagtgaac	13560
tgcagttgcc	atgttttacg	gcagtgagag	cagagatagc	gctgatgtcc	ggcggtgctt	13620
ttgccgttac	gcaccacccc	gtcagtagct	gaacaggagg	gacagctgat	agacacagaa	13680
gccactggag	cacctcaaaa	acaccatcat	acactaaatc	agtaagttgg	cagcatcacc	13740
cataattgtg	gtttcaaaat	cggctccgtc	gatactatgt	tatacgccaa	ctttgäaaac	13800
aactttgaaa	aagctgtttt	ctggtattta	aggttttaga	atgcaaggaa	cagtgaattg	13860
gagttcgtct	tgttataatt	agcttcttgg	ggtatcttta	aatactgtag	aaaagaggaa	13920
ggaaataata	aatggctaaa	atgagaatat	caccggaatt	gaaaaaactg	atcgaaaaat	13980
accgctgcgt	aaaagatacg	gaaggaatgt	ctcctgctaa	ggtatataag	ctggtgggag	14040
aaaatgaaaa	cctatattta	aaaatgacgg	acagccggta	taaagggacc	acctatgatg	14100
tggaacggga	aaaggacatg	atgctatggc	tggaaggaaa	gctgcctgtt	ccaaaggtcc	14160
tgcactttga	acggcatgat	ggctggagca	atctgctcat	gagtgaggcc	gatggcgtcc	14220
tttgctcgga	. agagtatgaa	gatgaacaaa	gccctgaaaa	gattatcgag	ctgtatgcgg	14280
agtgcatcag	gctctttcac	tccatcgaca	tatcggattg	tccctatacg	aatagcttag	14340

WO 2004/063359 PCT/EP2004/000099 166/357

acagccgctt	agccgaattg	gattacttac	tgaataacga	tctggccgat	gtggattgcg	14400
aaaactggga	agaagacact	ccatttaaag	atccgcgcga	gctgtatgat	tttttaaaga	14,460
cggaaaagcc	cgaagaggaa	cttgtctttt	cccacggcga	cctgggagac	agcaacatct	14520
ttgtgaaaga	tggcaaagta	agtggcttta	ttgatcttgg	gagaagcggc	agggcggaca	14580
agtggtatga	cattgccttc	tgcgtccggt	cgatcaggga	ggatatcggg	gaagaacagt	14640
atgtcgagct	attttttgac	ttactgggga	tcaagcctga	ttgggagaaa	ataaaatatt	14700
atattttact	ggatgaattg	ttttagtacc	tagatgtggc	gcaacgatgc	cggcgacaag	14760
caggagcgca	ccgacttctt	ccgcatcaag	tgttttggct	ctcaggccga	ggcccacggc	14820
aagtatttgg	gcaaggggtc	gctggtattc	gtgcagggca	agattcggaa	taccaagtac	14880
gagaaggacg	gccagacggt	ctacgggacc	gacttcattg	ccgataaggt	ggattatctg	14940
gacaccaagg	caccaggegg	gtcaaatcag	gaataagggc	acattgcccc	ggcgtgagtc	15000
ggggcaatcc	cgcaaggagg	gtgaatgaat	cggacgtttg	accggaaggc	atacaggcaa	15060
gaactgatcg	acgcggggtt	ttccgccgag	gatgccgaaa	ccatcgcaag	ccgcaccgtc	15120
atgcgtgcgc	cccgcgaaac	cttccagtcc	gtcggctcga	tggtccagca	agctacggcc	15180
aagatcgago	gegacagegt	gcaactggct	cccctgccc	tgcccgcgcc	ateggeegee	15240
gtggagcgtt	cgcgtcgtct	cgaacaggag	gcggcaggtt	: tggcgaagtc	gatgaccatc	15300
gacacgcgag	gaactatgac	gaccaagaag	r cgaaaaaccg	ccggcgagga	cctggcaaaa	15360
caggtcagcg	g aggecaagea	ggeegegttg	rctgaaacaca	ı cgaagcagca	gatcaaggaa	15420
atgcagettt	: ccttgttcga	a tattgcgccg	g tggccggaca	cgatgcgago	gatgccaaac	15480
gacacggcc	getetgeeet	gttcaccacg	g cgcaacaaga	a aaatcccgcg	g cgaggcgctg	15540
caaaacaagg	g tcattttcca	a cgtcaacaag	g gacgtgaaga	a tcacctacac	cggcgtcgag	15600

WO 2004/063359 PCT/EP2004/000099 167/357

ctgcgggccg	acgatgacga	actggtgtgg	cagcaggtgt	tggagtacgc	gaagcgcacc	15660
cctatcggcg	agccgatcac	cttcacgttc	tacgagcttt	gccaggacct	gggctggtcg	15720
atcaatggcc	ggtattacac	gaaggccgag	gaatgcctgt	cgcgcctaca	ggcgacggcg	15780
atgggcttca	cgtccgaccg	cgttgggcac	ctggaatcgg	tgtcgctgct	gcaccgcttc	15840
cgcgtcctgg	accgtggcaa	gaaaacgtcc	cgttgccagg	tcctgatcga	cgaggaaatc	15900
gtcgtgctgt	ttgctggcga	ccactacacg	aaattcatat	gggagaagta	ccgcaagctg	15960
tcgccgacgg	cccgacggat	gttcgactat	ttcagctcgc	accgggagcc	gtacccgctc	16020
aagctggaaa	ccttccgcct	catgtgcgga	tcggattcca	cccgcgtgaa	gaagtggcgc	16080
gagcaggtcg	gcgaagcctg	cgaagagttg	cgaggcagcg	gcctggtgga	acacgcctgg	16140
gtcaatgatg	acctggtgca	ttgcaaacgc	tagggccttg	tggggtcagt	teeggetggg	16200
ggttcagcag	ccagcgcttt	actggcattt	caggaacaag	cgggcactgc	tcgacgcact	16260
tgcttcgctc	agtatcgctc	gggacgcacg	gegegeteta	cgaactgccg	ataaacagag	16320
gattaaaatt	gacaattgtg	attaaggctc	agattcgacg	gcttggagcg	geegaegtge	16380
aggatttccg	cgagatccga	ttgtcggccc	tgaagaaagc	tccagagatg	ttcgggtccg	16440
tttacgagca	cgaggagaaa	aagcccatgg	aggcgttcgc	tgaacggttg	cgagatgccg	16500
tggcattcgg	cgcctacatc	gacggcgaga	tcattgggct	gtcggtcttc	aaacaggagg	16560
acggccccaa	ggacgctcac	aaggcgcatc	tgtccggcgt	tttcgtggag	cccgaacagc	16620
gaggccgagg	ggtcgccggt	atgctgctgc	gggcgttgcc	ggcgggttta	ttgctcgtga	16680
tgatcgtccg	acagattcca	acgggaatct	ggtggatgcg	catcttcatc	ctcggcgcac	16740
ttaatatttc	gctattctgg	agcttgttgt	ttatttcggt	ctaccgcctg	ccgggcgggg	16800
tegeggegae	ggtaggcgct	gtgcagccgc	tgatggtcgt	gttcatctct	gccgctctgc	16860

taggtagece gatacgattg atggeggtee tggggggetat ttgeggaact gegggegtgg 169) 20
cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 169	980
egggggeggt ttecatggeg tteggaaceg tgetgaeeeg caagtggeaa ceteeegtge 170	040
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 171	100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 173	160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 172	220
tegaacetae agttgtttee ttactggget tteteageee eagatetggg gtegateage 172	280
cggggatgca tcaggccgac agtcggaact tcgggtcccc gacctgtacc attcggtgag 17	340
caatggatag gggagttgat atcgtcaacg ttcacttcta aagaaatagc gccactcagc 17	400
ttcctcageg gctttatcca gegatttcct attatgtcgg catagttctc aagatcgaca 17	460
gcctgtcacg gttaagcgag aaatgaataa gaaggctgat aattcggatc tctgcgaggg 17	520
agatgatatt tgatcacagg cagcaacgct ctgtcatcgt tacaatcaac atgctaccct 17	580
ccgcgagate atccgtgttt caaacccggc agcttagttg ccgttettee gaatagcate 17	640
ggtaacatga gcaaagtetg eegeettaca aeggetetee egetgaegee gteeeggaet 17	7700
gatgggctgc ctgtatcgag tggtgatttt gtgccgagct gccggtcggg gagctgttgg 17	77 6 0
ctggctggtg gcaggatata ttgtggtgta aacaaattga cgcttagaca acttaataac 17	7820
acattgcgga cgtttttaat gtactggggt ggtttttctt ttcaccagtg agacgggcaa 17	7880
cagetgattg ecetteaceg eetggeeetg agagagttge ageaageggt eeaegetggt 17	7940
ttgccccagc aggcgaaaat cctgtttgat ggtggttccg aaatcggcaa aatcccttat 18	8000
aaatcaaaag aatagcccga gatagggttg agtgttgttc cagtttggaa caagagtcca 18	8060
ctattaaaga acgtggactc caacgtcaaa gggcgaaaaa ccgtctatca gggcgatggc 18	8120

ceactacgtg aaccatcacc caaatcaagt tttttggggt cgaggtgccg taaagcacta 18180

aatcggaacc ctaaagggag ceecegattt agagettgac ggggaaagge ggegaacgtg 18240

gcgagaaagg aagggaagaa agegaaagga gegggegeca tteaggetge geaactgttg 18300

ggaagggega teggtgeggg cetetteget attacgecag etggegaaag ggggatgtge 18360

tgcaaggega ttaagttggg taacgecagg gtttteccag teacgacgtt gtaaaacgac 18420

ggccagtgaa ttegageteg gtacceggg : 18449

<210> 42

<211> 17593

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 42

cegggetggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag

aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg

60

120

WO 2004/063359 PCT/EP2004/000099 170/357

aaaacttggc	cctcactgac	agatgagggg	cggacgttga	cacttgaggg	gccgactcac	180
ccggcgcggc	gttgacagat	gaggggcagg	ctcgatttcg	gccggcgacg	tggagctggc	240
cagootogca	aatcggcgaa	aacgcctgat	tttacgcgag	tttcccacag	atgatgtgga	300
caagcctggg	gataagtgcc	ctgcggtatt	gacacttgag	gggcgcgact	actgacagat	360
gaggggcgcg	atccttgaca	cttgaggggc	agagtgctga	cagatgaggg	gcgcacctat	420
tgacatttga	ggggctgtcc	acaggcagaa	aatccagcat	ttgcaagggt	ttccgcccgt	480
ttttcggcca	ccgctaacct	gtcttttaac	ctgcttttaa	accaatattt	ataaaccttg	540
tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcgcac	gccgaagggg	ggtgccccc	600
cttctcgaac	cctcccggcc	cgctaacgcg	ggçctcccat	cccccaggg	gctgcgcccc	660
teggeegega	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	cattgccggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	eggeggeetg	840
ggtggcggcc	tgcccttcac	tteggeegte	ggggcattca	cggacttcat	ggcggggccg	900
gcaattttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg	gacctttaca	gaattactct	atgaagcgcc	atatttaaaa	agctaccaag	1080
acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	aatactgata	1140
agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg	cgcttatcaa	tatatctata	gaatgggcaa	agcataaaaa	cttgcatgga	1260
ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380

WO 2004/063359 PCT/EP2004/000099 171/357

agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgccgctc	aattcgctgc	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cageggeeag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg	tcttccggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat	agccccactg	ttcgtccatt	tccgcgcaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg	aggttaccga	ctgcggcctg	agttttttaa	gtgacgtaaa	atcgtgttga	1800
ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg	agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc	cgtcgatact	atgttatacg	ccaactttga	aaacaacttt	gaaaaagctg	2160
ttttctggta	tttaaggttt	tagaatgcaa	ggaacagtga	attggagttc	gtcttgttat	2220
aattagcttc	ttggggtatc	tttaaatact	gtagaaaaga	ggaaggaaat	aataaatggc	2280
taaaatgaga	atatcaccgg	aattgaaaaa	actgatcgaa	aaataccgct	gcgtaaaaga	2340
tacggaagga	atgtctcctg	ctaaggtata	taagctggtg	ggagaaaatg	aaaacctata	2400
tttaaaaatg	acggacagcc	ggtataaagg	gaccacctat	gatgtggaac	gggaaaagga	2460
catgatgcta	tggctggaag	gaaagctgcc	tgttccaaag	gtcctgcact	ttgaacggca	2520
tgatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640

WO 2004/063359 PCT/EP2004/000099 172/357

tcactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac	ttactgaata	acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt	aaagatccgc	gcgagctgta	tgattttta	aagacggaaa	agcccgaaga	2820
ggaacttgtc	ttttcccacg	gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc	tttattgatc	ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc	cggtcgatca	gggaggatat	cggggaagaa	cagtatgtcg	agctattttt	3000
tgacttactg	gggatcaagc	ctgattggga	gaaaataaaa	tattatattt	tactggatga	3060
attgttttag	tacctagatg	tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tcttccgcat	caagtgtttt	ggctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt	attcgtgcag	ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg	gaccgacttc	attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa	tcaggaataa	gggcacattg	ccccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc	cgaggatgcc	gaaaccatcg	caagccgcac	cgtcatgcgt	gegeeeegeg	3480
aaaccttcca	gteegtegge	tcgatggtcc	agcaagctac	ggccaagatc	gagcgcgaca	3540
gcgtgcaact	ggeteecet	. gecetgeeeg	cgccatcggc	cgccgtggag	cgttcgcgtc	3600
gtctcgaaca	ı ggaggeggea	ggtttggcga	agtcgatgac	catcgacacg	cgaggaacta	3660
tgacgaccaa	ı gaagcgaaaa	accgccggcg	aggacctggc	aaaacaggto	agegaggeea	3720
agcaggccgc	gttgctgaaa	ı cacacgaago	agcagatcaa	ggaaatgcag	ctttccttgt	3780
tcgatattgo	geegtggeeg	g gacacgatgo	gagegatgee	aaacgacacg	gcccgctctg	3840
ccctgttcac	c cacgcgcaad	aagaaaatco	cgcgcgaggc	gctgcaaaa	: aaggtcattt	3900

WO 2004/063359 PCT/EP2004/000099 173/357

tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctate	ggcgagccga	4020
tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggc	cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg	gcacctggaa	teggtgtege	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200
gcaagaaaac	gtcccgttgc	caggtcctga	tcgacgagga	aatcgtcgtg	ctgtttgctg	4260
gcgaccacta	cacgaaattc	atatgggaga	agtaccgcaa	gctgtcgccg	acggcccgac	4320
ggatgttcga	ctatttcagc	tcgcaccggg	agccgtaccc	gctcaagctg	gaaaccttcc	4380
gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
cctgcgaaga	gttgcgaggc	ageggeetgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	gcagccagcg	4560
ctttactggc	atttcaggaa	caagcgggca	ctgctcgacg	cacttgcttc	gctcagtatc	4620
gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
tgtgattaag	gctcagattc	gacggcttgg	agcggccgac	gtgcaggatt	tccgcgagat	4740
ccgattgtcg	gccctgaaga	aagctccaga	gatgttcggg	teegtttaeg	agcacgagga	4800
gaaaaagccc	atggaggcgt	tcgctgaacg	gttgcgagat	gccgtggcat	teggegeeta	4860
catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagcgaggcc	gaggggtcgc	4980
cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	tccgacagat	5040
tccaacggga	atctggtgga	tgcgcatctt	catcctcggc	gcacttaata	tttcgctatt	5100
ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160

WO 2004/063359 PCT/EP2004/000099 174/357

cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg	5220
attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tggtgttgac	5280
accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat	5340
ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac	5400
cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc	5460
gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg	5520
agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt	5580
ttccttactg ggctttctca gccccagatc tggggtcgat cagccgggga tgcatcaggc	5640
cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt	5700
tgatategte aacgtteact tetaaagaaa tagegeeact cagetteete ageggettta	5760
tecagegatt tectattatg teggeatagt teteaagate gaeageetgt caeggttaag	5820
cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca	5880
caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt	5940
gtttcaaacc eggeagetta gttgeegtte tteegaatag categgtaac atgageaaag	6000
tctgccgcct tacaacggct ctcccgctga cgccgtcccg gactgatggg ctgcctgtat	6060
cgagtggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga	6120
tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt	6180
taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc	6240
accgcctggc cctgagagag ttgcagcaag cggtccacgc tggtttgccc cagcaggcga	6300
aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc	6360
ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg	6420

WO 2004/063359 PCT/EP2004/000099 175/357

actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480 6540 cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6600 ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga 6660 agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6720 cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6780 tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840 6900 cgaggagcct cgtcctgtca caactaccaa catggagtac gataagggcc agttccgcca 6960 gctcattaag agccagttca tgggcgttgg catgatggcc gtcatgcatc tgtacttcaa 7020 gtacaccaac gctcttctga tccagtcgat catccgctga aggcgctttc gaatctggtt 7080 aagatccacg tcttcgggaa gccagcgact ggtgacctcc agcgtccctt taaggctgcc aacagettte teageeaggg eeageeeaag acegacaagg eeteeeteea gaacgeegag 7140 aagaactgga ggggtggtgt caaggaggag taagctcctt attgaagtcg gaggacggag 7200 7260 cggtgtcaag aggatattct tcgactctgt attatagata agatgatgag gaattggagg tagcatagct tcatttggat ttgctttcca ggctgagact ctagcttgga gcatagaggg 7320 tcctttggct ttcaatattc tcaagtatct cgagtttgaa cttattccct gtgaaccttt 7380 7440 tattcaccaa tgagcattgg aatgaacatg aatctgagga ctgcaatcgc catgaggttt 7500 tcgaaataca tccggatgtc gaaggcttgg ggcacctgcg ttggttgaat ttagaacgtg 7560 gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7620 tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680

WO 2004/063359 PCT/EP2004/000099 176/357

gtaaggattt	eggcaegget	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaacatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ggtggagagc	ttataccgag	ctcccaaatc	7920
tgtccagatc	atggttgacc	ggtgcctgga	tcttcctata	gaatcatcct	tattcgttga	7980
	tctggagtga					8040
	gaaaaatgtg					8100
	agagacggac					8160
	cggctctgag					8220
	tggaaaggct					8280
	gcttcatcga					8340
	cggcgaaata					8400
	agattcacga					8460
						8520
	cctaattggc					8580
	cggtgtatga					
	agctggcagt					8640
	gtaggcagct					8700
	gtcagtccaa					8760
	ctetttettt					8820
tatttcccct	: aagtaagtad	: tttgctacat	: ccatactcca	teetteeeat	cccttattcc	8880
tttgaacctt	tcagttcgag	g ctttcccact	tcatcgcago	: ttgactaaca	gctaccccgc	8940

WO 2004/063359 PCT/EP2004/000099 177/357

ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agegteteeg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	ateggeactt	tgcatcggcc	gegetecega	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agegggtteg	gcccattcgg	9360
accgcaagga	atcggtcaat	acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	tcgcgcaggc	9480
tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	tcgtgcacgc	9540
ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	cggagcttgc	9720
aggategeeg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga	gccgggactg	tegggegtae	acaaatcgcc	cgcagaagcg	eggeegtetg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960
gagggcaaag	gaatagagta	gatgccgacc	gegggatega	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080
ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	. caacaagaat	aaaacgcgtt	ttcgggttta	10200

WO 2004/063359 PCT/EP2004/000099 178/357

cctcttccag	atacagctca	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagagc	tattttcatt	ttcgggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380
tccacgcgac	tatatatttg	tctctaattg	tactttgaca	tgctcctctt	ctttactctg	10440
atagcttgac	tatgaaaatt	ccgtcaccag	cncctgggtt	cgcaaagata	attgcatgtt	10500
tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
canttcctac	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactctc	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tetecaeteg	acctgcaggc	atgcaagctt	ttttcgagtt	10800
tttttttt	ttctttgtga	aggatttatt	gttattggta	tccattttt	attggaagac	10860
aagataagtt	aatattgatt	ttgcttaaag	attaaaagga	aatcagaaaa	cgacaataaa	10920
aaatgtaacg	gacaaactat	ggtgtcgatt	ataagtctaa	atccttaaaa	aatgacaacg	10980
agttgctttc	ctctgaaaac	aattcttttg	tctttgcaag	aaaggtttct	tttttgtttg	11040
cttgcattac	ttaaacatca	aatcaaatga	aaggaataaa	gcagatttga	gggcgaataa	11100
ggattttctg	gtcaacaaga	tgtgagtgac	acctaaggaa	ctaaatgcca	ttcatttgtt	11160
ttaaaacgac	atcaaagatt	gatgatcaac	aggattgaga	gagagaaaaa	gaactcgtgt	11220
catttatttc	tgttgactga	aattttatat	ttagaaaaaa	tgtcaaatct	atagctttag	11280
ctatattaca	taacatttga	aataataata	ataaaaaaag	acacattaga	gacacttttc	11340
aaactctaaa	taactgtcta	taaacacaaa	gaaaacaaag	acctctataa	. caacttatta	11400
gatttttctc	gtacttttgt	ctaaagatga	tgtattcttg	ttatcccaca	cttctttcat	11460

WO 2004/063359 PCT/EP2004/000099 179/357

ttgttcttga	tgctactaaa	tatacaaaat	ttcttttttg	caagagatat	tattccaaaa	11520
attttcaaaa	agaaattttt	ttcacaatag	cagttgatcg	tgtaacccaa	agaggttett	11580
tgttattttg	cacttccgct	ttgcggtgat	gcatattcaa	agtaatatat	ggaataaaca	11640
acgtgtttaa	gcatgaaaga	aaggaaacaa	aggccgcttt	gaacaaatgc	ataatatttc	11700
agacaaaaat	: gatctaaagc	aagcagtaaa	tcaaacaaga	aacattgctg	attcgcgtta	11760
gaaaacgata	a aaagtctaat	aagccactaa	gtatacttca	atgaactttt	tgtatgctta	11820
tggtccaato	agaccaataa	tttgtgacca	ttcctgaggt	ggctttggtg	atgcggaaac	11880
agaaaaaaat	tttctcacca	atcgatttaa	aaaacaattt	ctgctttgaa	ccaaaacttt	11940
ttttttctct	ttaatcatta	actttatcaa	gtatgtacct	accctcaaag	tcctcactca	12000
agcacaatta	a tgctaacatt	gttccacctt	ctctttagaa	atgttgtgga	tttggaatgc	12060
cctgatcgt	t ttcgttaccg	tgattggcat	ggaagtgatt	gctgcactgg	cacacaaata	12120
catcatgca	c ggctggggtt	ggggatggca	tctttcacat	catgaaccgc	gtaaaggtgc	12180
gtttgaagt	t aacgatcttt	atgccgtggt	ttttgctgca	ttatcgatcc	tgctgattta	12240
tctgggcag	t acaggaatgt	ggccgctcca	gtggattggc	gcaggtatga	cggcgtatgg	12300
attactcta	t tttatggtgc	acgacgggct	ggtgcatcaa	cgttggccat	tccgctatat	12360
tccacgcaa	g ggctacctca	aacggttgta	tatggcgcac	cgtatgcatc	acgccgtcag	12420
gggcaaaga	a ggttgtgtt	cttttggctt	cctctatgcg	cegeceetgt	caaaacttca	12480
ggcgacgct	c cgggaaagac	atggcgctag	agcgggcgct	gccagagatg	cgcagggcgg	12540
ggaggatga	g cccgcatccg	ggaagtaagg	geetgaeeag	aggcggccag	cagcagcgtt	12600
aatttttcg	g gegtggtegt	: tgactgccgc	tgatcccaaa	gcttggcgta	atcatggtca	12660
tagctgttt	c ctgtgtgaaa	ttgttatccg	ctcacaatto	cacacaacat	acgagccgga	12720

agcataaagt gtaaagcctg gggtgcctaa tgagtgagct aactcacatt aattgcgttg 12780 cgctcactgc ccgctttcca gtcgggaaac ctgtcgtgcc agctgcatta atgaatcggc 12840 caacgcgcgg ggagaggcgg tttgcgtatt gggccaaaga caaaagggcg acattcaacc 12900 gattgaggga gggaaggtaa atattgacgg aaattattca ttaaaggtga attatcaccg 12960 tcaccgactt gagccatttg ggaattagag ccagcaaaat caccagtagc accattacca 13020 ttagcaaggc cggaaacgtc accaatgaaa ccatcgatag cagcaccgta atcagtagcg 13080 acagaatcaa gtttgccttt agcgtcagac tgtagcgcgt tttcatcggc attttcggtc 13140 atageceet tattagegtt tgecatettt teataateaa aateaeegga aeeagageea 13200 ccaccggaac cgcctccctc agagccgcca ccctcagaac cgccaccctc agagccacca 13260 ccctcagagc cgccaccaga accaccacca gagccgccgc cagcattgac aggaggcccg 13320 atctagtaac atagatgaca ccgcgcgcga taatttatcc tagtttgcgc gctatatttt 13380 gttttctatc gcgtattaaa tgtataattg cgggactcta atcataaaaa cccatctcat 13440 aaataacgtc atgcattaca tgttaattat tacatgctta acgtaattca acagaaatta 13500 tatgataatc atcgcaagac cggcaacagg attcaatctt aagaaacttt attgccaaat 13560 gtttgaacga tcggggatca tccgggtctg tggcgggaac tccacgaaaa tatccgaacg 13620 cagcaagata tegeggtgca teteggtett geetgggeag tegeeggega egeegttgat 13680 gtggacgccg ggcccgatca tattgtcgct caggatcgtg gcgttgtgct tgtcggccgt 13740 tgctgtcgta atgatatcgg caccttcgac cgcctgttcc gcagagatcc cgtgggcgaa 13800 gaactecage atgagatece egegetggag gateatecag eeggegteee ggaaaaegat 13860 tccgaagccc aacctttcat agaaggcggc ggtggaatcg aaatctcgtg atggcaggtt 13920 gggcgtcgct tggtcggtca tttcgaaccc cagagtcccg ctcagaagaa ctcgtcaaga 13980

WO 2004/063359 PCT/EP2004/000099 181/357

aggcgataga	aggcgatgcg	ctgcgaatcg	ggagcggcga	taccgtaaag	cacgaggaag	14040
cggtcagccc	attcgccgcc	aagctcttca	gcaatatcac	gggtagccaa	cgctatgtcc	14100
tgatagcggt	ccgccacacc	cageeggeea	cagtcgatga	atccagaaaa	gcggccattt	14160
tccaccatga	tattcggcaa	gcaggcatcg	ccatgggtca	cgacgagatc	atcgccgtcg	14220
ggcatgcgcg	ccttgagcct	ggcgaacagt	teggetggeg	cgagcccctg	atgctcttcg	14280
tccagatcat	cctgatcgac	aagaccggct	tccatccgag	tacgtgctcg	ctcgatgcga	14340
tgtttcgctt	ggtggtcgaa	tgggcaggta	gccggatcaa	gcgtatgcag	ccgccgcatt	14400
gcatcagcca	tgatggatac	tttctcggca	ggagcaaggt	gagatgacag	gagatcctgc	14460
cccggcactt	cgcccaatag .	cagccagtcc	cttcccgctt	cagtgacaac	gtcgagcaca	14520
gctgcgcaag	gaacgcccgt	cgtggccagc	cacgatagcc	gegetgeete	gtcctgcagt	14580
tcattcaggg	caccggacag	gtcggtcttg	acaaaaagaa	ccgggcgccc	ctgcgctgac	14640
agccggaaca	cggcggcatc	agagcagccg	attgtctgtt	gtgcccagtc	atagccgaat	14700
agecteteca	cccaagcggc	cggagaacct	gcgtgcaatc	catcttgttc	aatcatgcga	14760
aacgatccag	atccggtgca	gattatttgg	attgagagtg	aatatgagac	tctaattgga	14820
taccgagggg	aatttatgga	acgtcagtgg	agcatttttg	acaagaaata	tttgctagct	14880
gatagtgacc	ttaggcgact	tttgaacgcg	caataatggt	ttctgacgta	tgtgcttagc	14940
tcattaaact	ccagaaaccc	gcggctgagt	ggctccttca	acgttgcggt	tctgtcagtt	15000
ccaaacgtaa	aacggcttgt	cccgcgtcat	cggcgggggt	cataacgtga	ctcccttaat	15060
tctccgctca	tgatcagatt	gtcgtttccc	gccttcagtt	taaactatca	gtgtttgaca	15120
ggatatattg	gcgggtaaac	ctaagagaaa	agagcgttta	ttagaataat	cggatattta	15180
aaagggcgtg	aaaaggttta	teegttegte	catttgtatg	tgcatgccaa	ccacagggtt	15240

WO 2004/063359 PCT/EP2004/000099 182/357

ccccagatct	ggegeeggee	agcgagacga	gcaagattgg	ccgccgcccg	aaacgatccg	15300
acagegegee	cagcacaggt	gcgcaggcaa	attgcaccaa	cgcatacagc	gccagcagaa	15360
tgccatagtg	ggcggtgacg	tcgttcgagt	gaaccagatc	gcgcaggagg	cccggcagca	15420
ccggcataat	caggccgatg	ccgacagcgt	cgagcgcgac	agtgctcaga	attacgatca	15480
ggggtatgtt	gggtttcacg	tetggeetee	ggaccagcct	ccgctggtcc	gattgaacgc	15540
gcggattctt	tatcactgat	aagttggtgg	acatattatg	tttatcagtg	ataaagtgtc	15600
aagcatgaca	aagttgcagc	cgaatacagt	gatccgtgcc	gccctggacc	tgttgaacga	15660
ggtcggcgta	gacggtctga	cgacacgcaa	actggcggaa	cggttggggg	ttcagcagcc	15720
ggcgctttac	tggcacttca	ggaacaagcg	ggcgctgctc	gacgcactgg	ccgaagccat	15780
gctggcggag	aatcatacgc	attcggtgcc	gagagccgac	gacgactggc	gctcatttct	15840
gatcgggaat	gcccgcagct	tcaggcaggc	gctgctcgcc	taccgcgatg	gcgcgcgcat	15900
ccatgccggc	acgcgaccgg	gcgcaccgca	gatggaaacg	gccgacgcgc	agcttcgctt	15960
cctctgcgag	gcgggttttt	cggccgggga	cgccgtcaat	gcgctgatga	caatcagcta	16020
cttcactgtt	ggggccgtgc	ttgaggagca	ggccggcgac	agcgatgccg	gegagegegg	16080
cggcaccgtt	gaacaggctc	cgctctcgcc	gctgttgcgg	gccgcgatag	acgccttcga	16140
cgaagccggt	ccggacgcag	cgttcgagca	gggactcgcg	gtgattgtcg	atggattggc	16200
gaaaaggagg	ctcgttgtca	ggaacgttga	aggaccgaga	aagggtgacg	attgatcagg	16260
accgctgccg	gagcgcaacc	cactcactac	agcagagcca	tgtagacaac	atcccctccc	16320
cctttccacc	gcgtcagacg	cccgtagcag	cccgctacgg	gctttttcat	gecetgecet	16380
agcgtccaag	cctcacggcc	gcgctcggcc	tetetggegg	ccttctggcg	ctcttccgct	16440
tectegetea	ctgactcgct	gcgctcggtc	gttcggctgc	ggcgagcggt	atcagctcac	16500

tcaaaggcgg	taatacggtt	atccacagaa	tcaggggata	acgcaggaaa	gaacatgtga	16560
gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	cgttgctggc	gtttttccat	16620
aggctccgcc	cccctgacga	gcatcacaaa	aatcgacgct	caagtcagag	gtggcgaaac	16680
ccgacaggac	tataaagata	ccaggcgttt	cccctggaa	gctccctcgt	gcgctctcct	16740
gttccgaccc	tgccgcttac	cggatacctg	teegeettte	tcccttcggg	aagcgtggcg	16800
cttttccgct	gcataaccct	gcttcggggt	cattatagcg	attttttcgg	tatatccatc	16860
ctttttcgca	cgatatacag	gattttgcca	aagggttcgt	gtagactttc	cttggtgtat	16920
ccaacggcgt	cageegggea	ggataggtga	agtaggccca	cccgcgagcg	ggtgttcctt	16980
cttcactgtc	ccttattcgc	acctggcggt	gctcaacggg	aatcctgctc	tgcgaggctg	17040
geeggetace	gccggcgtaa	cagatgaggg	caagcggatg	gctgatgaaa	ccaagccaac	17100
caggaagggc	agcccaccta	tcaaggtgta	ctgccttcca	gacgaacgaa	gagcgattga	17160
ggaaaaggcg	gcggcggccg	gcatgagcct	gtcggcctac	ctgctggccg	tcggccaggg	17220
ctacaaaatc	acgggcgtcg	tggactatga	gcacgtccgc	gagctggccc	gcatcaatgg	17280
cgacctgggc	cgcctgggcg	gcctgctgaa	actctggctc	accgacgacc	cgcgcacggc	17340
geggtteggt	gatgccacga	tcctcgccct	gctggcgaag	atcgaagaga	agcaggacga	17400
gcttggcaag	gtcatgatgg	gcgtggtccg	cccgagggca	gagccatgac	ttttttagcc	17460
gctaaaacgg	ccggggggtg	cgcgtgattg	ccaagcacgt	ccccatgcgc	tccatcaaga	17520
agagcgactt	cgcggagctg	gtgaagtaca	tcaccgacga	gcaaggcaag	accgagcgcc	17580
tttgcgacgc	tca					17593

<210> 43

<211> 16954

184/357

<212> DNA <213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<400> 43

ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 120 180 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240 cagcetegea aateggegaa aacgeetgat tttacgegag tttcccacag atgatgtgga 300 360 caagectggg gataagtgee etgeggtatt gacaettgag gggegegaet aetgacagat gaggggegeg atcettgaca ettgagggge agagtgetga eagatgaggg gegeacetat 420 tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480 540 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 600 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc

WO 2004/063359 PCT/EP2004/000099 185/357

cttctcgaac	cctcccggcc	cgctaacgcg	ggcctcccat	cccccaggg	gctgcgcccc	660
teggeegega	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	cattgccggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	cggcggcctg	840
ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaatttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg	gacctttaca	gaattactct	atgaagegee	atatttaaaa	agctaccaag	1080
acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	aatactgata	1140
agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg	cgcttatcaa	tatatctata	gaatgggcaa	agcataaaaa	cttgcatgga	1260
ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgccgctc	aattcgctgc	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg	f tetteeggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat	agccccactg	ttcgtccatt	teegegeaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg	g aggttaccga	ctgcggcctg	agtttttaa	ı gtgacgtaaa	atcgtgttga	1800
ggccaacgc	c cataatgcgg	gctgttgccc	ggcatccaac	gecatteate	gccatatcaa	1860

WO 2004/063359 PCT/EP2004/000099 186/357

tgattttctg gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt to	gccatgttt	1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg tt	taegcacca	1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact gg	gagcacctc	2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat t	gtggtttca	2100
aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt g	aaaaagctg	2160
ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagttc g	tcttgttat	2220
aattagcttc ttggggtatc tttaaatact gtagaaaaga ggaaggaaat a	ataaatggc	2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct g	rcgtaaaaga	2340
tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg a	aaacctata	2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac g	ggaaaagga	2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact t	tgaacggca	2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct c	eggaagagta	2580
tgaagatgaa caaageeetg aaaagattat egagetgtat geggagtgea t	tcaggctctt	2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc g	gettageega	2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact g	gggaagaaga	2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa a	agcccgaaga	2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga	aagatggcaa	2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggt	atgacattgc	2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg	agctattttt	3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt	tactggatga	3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag	cgcaccgact	3120

WO 2004/063359 PCT/EP2004/000099 187/357

tcttccgcat caagtgtttt gg	ctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt attcgtgcag gg	caagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg gaccgacttc at	tgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa tcaggaataa gg	gcacattg	ccccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat gaatcggacg tt	tgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc cgaggatgcc ga	.aaccatcg	caagccgcac	cgtcatgcgt	gegeeeegeg	3480
aaaccttcca gtccgtcggc tc	gatggtcc	agcaagctac	ggccaagatc	gagcgcgaca	3540
gegtgeaact ggeteeect ge	cctgcccg	cgccatcggc	cgccgtggag	cgttcgcgtc	3.600
gtctcgaaca ggaggcggca gg	ıtttggcga	agtcgatgac	catcgacacg	cgaggaacta	3660
tgacgaccaa gaagcgaaaa ac	cgccggcg	aggacctggc	aaaacaggtc	agcgaggcca	3720
agcaggccgc gttgctgaaa ca	acacgaagc	agcagatcaa	ggaaatgcag	ctttccttgt	3780
tcgatattgc gccgtggccg ga	acacgatgc	gagcgatgcc	aaacgacacg	gcccgctctg	3840
ccctgttcac cacgcgcaac aa	agaaaatcc	cgcgcgaggc	gctgcaaaac	aaggtcattt	3900
tccacgtcaa caaggacgtg aa	agatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt gtggcagcag g	tgttggagt	acgcgaagcg	cacccctatc	ggcgagccga	4020
tcaccttcac gttctacgag c	tttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggc cgaggaatgc c	tgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg gcacctggaa t	cggtgtcgc	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200
gcaagaaaac gtcccgttgc c	aggtcctga	tcgacgagga	aatcgtcgtg	ctgtttgctg	4260
gcgaccacta cacgaaattc a	tatgggaga	agtaccgcaa	gctgtcgccg	g acggcccgac	4320
ggatgttcga ctatttcagc t	cgcaccggg	agccgtaccc	: gctcaagcto	g gaaaccttcc	4380

WO 2004/063359 PCT/EP2004/000099 188/357

gcctcatgtg	cggatcggat	tecaceegeg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
cctgcgaaga	gttgcgaggc	ageggeetgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	gcagccagcg	4560
ctttactggc	atttcaggaa	caagcgggca	ctgctcgacg	cacttgcttc	gctcagtatc	4620
gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
tgtgattaag	gctcagattc	gacggcttgg	agcggccgac	gtgcaggatt	tccgcgagat	4740
ccgattgtcg	gccctgaaga	aagctccaga	gatgttcggg	tccgtttacg	agcacgagga	4800
gaaaaagccc	atggaggcgt	tcgctgaacg	gttgcgagat	gccgtggcat	teggegeeta	4860
catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagcgaggcc	gaggggtcgc	4980
cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	tccgacagat	5040
tccaacggga	atctggtgga	tgcgcatctt	catectegge	gcacttaata	tttcgctatt	5100
ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160
cgctgtgcag	ccgctgatgg	tcgtgttcat	ctctgccgct	ctgctaggta	gcccgatacg	5220
attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	tggtgttgac	5280
accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	cggtttccat	5340
ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	tcacctttac	5400
cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	tgtttgatcc	5460
gccaatcccg	atgcctacag	gaaccaatgt	tctcggcctg	gcgtggctcg	gcctgatcgg	5520
agcgggttta	acctacttcc	tttggttccg	ggggateteg	cgactcgaac	ctacagttgt	5580
ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	tgcatcaggc	5640

WO 2004/063359 PCT/EP2004/000099 189/357

cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	ataggggagt	5700
tgatatcgtc	aacgttcact	tctaaagaaa	tagcgccact	cagcttcctc	agcggcttta	5760
tccagcgatt	tcctattatg	tcggcatagt	tctcaagatc	gacagcctgt	cacggttaag	5820
cgagaaatga	ataagaaggc	tgataattcg	gatctctgcg	agggagatga	tatttgatca	5880
caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	gatcatccgt	5940
gtttcaaacc	cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	atgagcaaag	6000
tctgccgcct	tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	ctgcctgtat	6060
cgagtggtga	ttttgtgccg	agctgccggt	cggggagctg	ttggctggct	ggtggcagga	6120
tatattgtgg	tgtaaacaaa	ttgacgctta	gacaacttaa	taacacattg	cggacgtttt	6180
taatgtactg	gggtggtttt	tcttttcacc	agtgagacgg	gcaacagctg	attgcccttc	6240
accgcctggc	cctgagagag	ttgcagcaag	cggtccacgc	tggtttgccc	cagcaggcga	6300
aaatcctgtt	tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	6360
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480
cacccaaatc	aagtttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	6540
ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600
agaaagcgaa	. aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt	cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720
tgggtaacgo	cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780
ctcggtaccc	ggggatettt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840
cgaggagcct	: cgtcctgtca	caactaccaa	. catggagtac	gataagggco	agttccgcca	6900

WO 2004/063359 PCT/EP2004/000099 190/357

gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac	gctcttctga	tccagtcgat	catccgctga	aggcgctttc	gaatctggtt	7020
aagatccacg	tcttcgggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140
aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag	aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320
teetttgget	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
tcgaaataca	tccggatgtc	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560
tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagco	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgactto	agcaacatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ı ggtggagagc	: țtataccgag	r ctcccaaatc	7920
tgtccagato	e atggttgaco	ggtgcctgga	a tcttcctata	gaatcatcct	tattcgttga	7980
cctagctgat	: tctggagtga	a cccagagggt	catgacttga	a gootaaaato	c cgccgcctcc	8040
accatttgta	a gaaaaatgto	g acgaactcg	t gagctctgta	a cagtgaccgg	g tgactctttc	8100
tggcatgcgg	g agagacggad	c ggacgcaga	g agaagggct	g agtaataag	c cactggccag	8160

WO 2004/063359 PCT/EP2004/000099 191/357

acagctctgg	cggctctgag	gtgcagtgga	tgattattaa	teegggaceg	gccgcccctc	8220
cgccccgaag	tggaaaggct	ggtgtgcccc	tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatagccgt	cggcgaaata	gcatgccatt	aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	agattcacga	gatagtacct	teteegaagt	aggtagagcg	agtacccggc	8460
gcgtaagctc	cctaattggc	ccatccggca	tctgtagggc	gtccaaatat	cgtgcctctc	8520
ctgctttgcc	cggtgtatga	aaccggaaag	gccgctcagg	agctggccag	cggcgcagac	8580
cgggaacaca	agctggcagt	cgacccatcc	ggtgctctgc	actcgacctg	ctgaggtccc	8640
tcagtccctg	gtaggcagct	ttgccccgtc	tgtccgcccg	gtgtgtcggc	ggggttgaca	8700
aggtcgttgc	gtcagtccaa	catttgttgc	catattttcc	tgctctcccc	accagctgct	8760
cttttcttt	ctctttcttt	teccatette	agtatattca	tcttcccatc	caagaacctt	8820
tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agegggtteg	gcccattcgg	9360
accgcaagga	atcggtcaat	. acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420

ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc	9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc	9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg	9600
gagcgaggcg atgttcgggg attcccaata cgaggtcgcc aacatcttct tctggaggcc	9660
gtggttggct tgtatggagc agcagacgcg ctacttcgag cggaggcatc cggagcttgc	9720
aggategeeg eggeteeggg egtatatget eegeattggt ettgaceaac tetateagag	9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt	9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg	9900
gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca gcactcgtcc	9960
gagggcaaag gaatagagta gatgeegaee gegggatega teeaettaae gttaetgaaa	10020
tcatcaaaca gettgacgaa tetggatata agategttgg tgtegatgte ageteeggag	10080
ttgagacaaa tggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga	10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta	10200
cctcttccag atacagetea tetgcaatge attaatgeat tgaetgeaac etagtaaege	10260
cttncagget ceggegaaga gaagaatage ttageagage tatttteatt ttegggagae	10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc	10380
tccacgcgac tatatatttg tctctaattg tactttgaca tgctcctctt ctttactctg	10440
atagettgae tatgaaaatt eegteaceag eneetgggtt egeaaagata attgeatgtt	10500
tcttccttga actctcaagc ctacaggaca cacattcatc gtaggtataa acctcgaaat	10560
canttectae taagatggta tacaatagta accatgeatg gttgeetagt gaatgeteeg	10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttacccaga	10680

WO 2004/063359 PCT/EP2004/000099 193/357

atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tetecaeteg	acctgcaggc	atgcaagctt	gagattaaaa	10800
tagataagga	aaagaaagtg	aaaagaaatt	cggaagcatg	gcacattett	ctttttataa	10860
atacatgcct	gactttcttt	ttccatcgat	atgatatatg	catatgatag	atatacaagc	10920
aatcttcttc	aaggagtttg	aaattttgtc	ctccaggagc	aaaaaaagt	tttttttat	10980
acatgtttgt	acacaagaat	agttaccaat	ttgctttggt	cttacgtgct	gcaagtttat	11040
atcgttttca	atttctttgt	ctttacattt	tctttgtcct	ttatctttcc	tcatttagtc	11100
tttgggagaa	ttaggaaaag	ggagcggaaa	ggtaagaaat	gcttgcgtat	tttactaatt	11160
cggcaaacat	ccaatttggc	aaacagcagc	ctgtgcaacg	ctctcgagat	gacagtatct	11220
ttgattacac	tctaaatctc	gatgacccga	ccaaaaagag	cgaacaaaga	aataatcttg	11280
tgcattcgaa	tatgatggaa	gattttttcc	cccttattct	aaatgttgac	atagcgtgta	11340
tgttatataa	acaaaaagaa	attgtacaaa	ctttctttc	ttctctttt	attttatctc	11400
tatgttgtgg	atttggaatg	ccctgatcgt	tttcgttacc	gtgattggca	tggaagtgat	11460
tgctgcactg	gcacacaaat	acatcatgca	cggctggggt	tggggatggc	atctttcaca	11520
tcatgaaccg	cgtaaaggtg	cgtttgaagt	taacgatctt	tatgccgtgg	tttttgctgc	11580
attatcgatc	ctgctgattt	atctgggcag	tacaggaatg	tggccgctcc	agtggattgg	11640
cgcaggtatg	acggcgtatg	gattactcta	ttttatggtg	cacgacgggc	tggtgcatca	11700
acgttggcca	ttccgctata	ttccacgcaa	gggctacctc	aaacggttgt	atatggcgca	11760
ccgtatgcat	cacgccgtca	ggggcaaaga	aggttgtgtt	tcttttggct	tcctctatgc	11820
geegeeeetg	tcaaaacttc	aggcgacgct	ccgggaaaga	catggcgcta	gagcgggcgc	11880
tgccagagat	gcgcagggcg	gggaggatga	gcccgcatcc	gggaagtaag	ggcctgacca	11940

WO 2004/063359 PCT/EP2004/000099 194/357

gaggcggcca	gcagcagcgt	taatttttcg	ggcgtggtcg	ttgactgccg	ctgatcccaa	12000
agcttggcgt	aatcatggtc	atagctgttt	cctgtgtgaa	attgttatcc	gctcacaatt	12060
ccacacaaca	tacgagccgg	aagcataaag	tgtaaagcct	ggggtgccta	atgagtgagc	12120
taactcacat	taattgcgtt	gcgctcactg	cccgctttcc	agtcgggaaa	cctgtcgtgc	12180
cagetgeatt	aatgaatcgg	ccaacgcgcg	gggagaggcg	gtttgcgtat	tgggccaaag	12240
acaaaagggc	gacattcaac	cgattgaggg	agggaaggta	aatattgacg	gaaattattc	12300
attaaaggtg	aattatcacc	gtcaccgact	tgagccattt	gggaattaga	gccagcaaaa	12360
tcaccagtag	caccattacc	attagċaagg	ccggaaacgt	caccaatgaa	accatcgata	12420
gcagcaccgt	aatcagtagc	gacagaatca	agtttgcctt	tagcgtcaga	ctgtagcgcg	12480
ttttcatcgg	cattttcggt	catagccccc	ttattagcgt	ttgccatctt	ttcataatca	12540
aaatcaccgg	aaccagagcc	accaccggaa	ccgcctccct	cagagccgcc	acceteagaa	12600
ccgccaccct	cagagccacc	accctcagag	ccgccaccag	aaccaccacc	agagccgccg	12660
ccagcattga	. caggaggccc	gatctagtaa	catagatgac	accgcgcgcg	ataatttatc	12720
ctagtttgcg	cgctatattt	tgttttctat	cgcgtattaa	atgtataatt	gcgggactct	12780
aatcataaaa	acccatctca	taaataacgt	catgcattac	atgttaatta	ttacatgctt	12840
aacgtaatto	: aacagaaatt	atatgataat	catcgcaaga	. ccggcaacag	gattcaatct	12900
taagaaactt	: tattgccaaa	tgtttgaacg	atcggggatc	atcegggtet	. gtggcgggaa	12960
ctccacgaaa	atateegaac	: gcagcaagat	atcgcggtgc	atctcggtct	tgcctgggca	13020
gtegeegeeg	g acgccgttga	tgtggacgco	gggcccgato	: atattgtcgc	: tcaggatcgt	13080
ggcgttgtg	c ttgtcggccg	ı ttgctgtcgt	: aatgatatcç	g gcacettega	ccgcctgttc	13140
cgcagagato	c ccgtgggcga	a agaactccag	g catgagatco	c ccgcgctgga	a ggatcatcca	13200

WO 2004/063359 PCT/EP2004/000099 195/357

gccggcgtcc	cggaaaacga	ttccgaagcc	caacctttca	tagaaggcgg	cggtggaatc	13260
gaaatctcgt	gatggcaggt	tgggcgtcgc	ttggtcggtc	atttcgaacc	ccagagtccc	13320
gctcagaaga	actcgtcaag	aaggcgatag	aaggcgatgc	gctgcgaatc	gggagcggcg	13380
ataccgtaaa	gcacgaggaa	gcggtcagcc	cattegeege	caagctcttc	agcaatatca	13440
cgggtagcca	acgctatgtc	ctgatagcgg	tccgccacac	ccagccggcc	acagtcgatg	13500
aatccagaaa	agcggccatt	ttccaccatg	atattcggca	agcaggcatc	gccatgggtc	13560
acgacgagat	categeegte	gggcatgcgc	gccttgagcc	tggcgaacag	tteggetgge	13620
gcgagcccct	gatgctcttc	gtccagatca	tcctgatcga	caagaccggc	ttccatccga	13680
gtacgtgctc	gctcgatgcg	atgtttcgct	tggtggtcga	atgggcaggt	agccggatca	13740
agcgtatgca	gccgccgcat	tgcatcagcc	atgatggata	ctttctcggc	aggagcaagg	13800
tgagatgaca	ggagatcctg	ccccggcact	tegeceaata	gcagccagtc	cetteceget	13860
tcagtgacaa	cgtcgagcac	agctgcgcaa	ggaacgcccg	tcgtggccag	ccacgatagc	13920
cgcgctgcct	cgtcctgcag	ttcattcagg	gcaccggaca	ggtcggtctt	gacaaaaaga	13980
accgggcgcc	cctgcgctga	cagccggaac	acggcggcat	cagagcagcc	gattgtctgt	14040
tgtgcccagt	catageegaa	tageetetee	acccaagcgg	ccggagaacc	tgcgtgcaat	14100
ccatcttgtt	caatcatgcg	aaacgatcca	gatccggtgc	agattatttg	gattgagagt	14160
gaatatgaga	ı ctctaattgg	ataccgaggg	gaatttatgg	aacgtcagtg	gagcattttt	14220
gacaagaaat	atttgctago	: tgatagtgac	: cttaggcgac	ttttgaacgo	gcaataatgg	14280
tttctgacgt	atgtgcttag	r ctcattaaac	: tccagaaacc	cgcggctgag	g tggctccttc	14340
aacgttgcgg	g ttctgtcagt	tccaaacgta	aaacggcttg	tecegegtea	tcggcggggg	14400
tcataacgt	g actcccttaa	ttctccgcto	c atgatcagat	tgtcgtttco	c cgccttcagt	14460

WO 2004/063359 PCT/EP2004/000099 196/357

ttaaactatc	agtgtttgac	aggatatatt	ggcgggtaaa	cctaagagaa	aagagcgttt	14520
attagaataa	tcggatattt	aaaagggcgt	gaaaaggttt	atccgttcgt	ccatttgtat	14580
gtgcatgcca	accacagggt	tccccagatc	tggcgccggc	cagcgagacg	agcaagattg	14640
gccgccgccc	gaaacgatcc	gacagcgcgc	ccagcacagg	tgcgcaggca	aattgcacca	14700
acgcatacag	cgccagcaga	atgccatagt	gggcggtgac	gtcgttcgag	tgaaccagat	14760
cgcgcaggag	gcccggcagc	accggcataa	tcaggccgat	gccgacagcg	tcgagcgcga	14820
cagtgctcag	aattacgatc	aggggtatgt	tgggtttcac	gtctggcctc	cggaccagcc	14880
tccgctggtc	cgattgaacg	cgcggattct	ttatcactga	taagttggtg	gacatattat	14940
gtttatcagt	gataaagtgt	caagcatgac	aaagttgcag	ccgaatacag	tgatccgtgc	15000
cgccctggac	ctgttgaacg	aggtcggcgt	agacggtctg	acgacacgca	aactggcgga	15060
acggttgggg	gttcagcagc	cggcgcttta	ctggcacttc	aggaacaagc	gggcgctgct	15120
cgacgcactg	gccgaagcca	tgctggcgga	gaatcatacg	catteggtge	cgagagccga	15180
cgacgactgg	cgctcatttc	tgatcgggaa	tgcccgcagc	ttcaggcagg	cgctgctcgc	15240
ctaccgcgat	ggcgcgcgca	tccatgccgg	cacgcgaccg	ggcgcaccgc	agatggaaac	15300
ggccgacgcg	cagcttcgct	tectetgega	ggcgggtttt	tcggccgggg	acgccgtcaa	15360
tgcgctgatg	acaatcagct	acttcactgt	tggggccgtg	cttgaggagc	aggccggcga	15420
cagcgatgcc	ggcgagcgcg	geggeacegt	tgaacaggct	cegetetege	cgctgttgcg	15480
ggccgcgata	gacgccttcg	acgaagccgg	teeggaegea	gcgttcgagc	agggactcgc	15540
ggtgattgtc	gatggattgg	cgaaaaggag	gctcgttgtc	aggaacgttg	aaggaccgag	15600
aaagggtgac	gattgatcag	gacegetgee	ggagcgcaac	ccactcacta	cagcagagcc	15660
atgtagacaa	cateceetee	ceetttccae	cgcgtcagac	gcccgtagca	geeegetaeg	15720

WO 2004/063359 PCT/EP2004/000099 197/357

ggctttttca	tgccctgccc	tagcgtccaa	gectcaegge	cgcgctcggc	ctctctggcg	15780
gccttctggc	gctcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	15840
cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggt	tatccacaga	atcaggggat	15900
aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	15960
gcgttgctgg	cgtttttcca	taggctccgc	cccctgacg	agcatcacaa	aaatcgacgc	16020
tcaagtcaga	ggtggcgaaa	cccgacagga	ctataaagat	accaggcgtt	tccccctgga	16080
agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggatacct	gtccgccttt	16140
ctcccttcgg	gaagcgtggc	gcttttccgc	tgcataaccc	tgcttcgggg	tcattatagc	16200
gattttttcg	gtatatccat	cctttttcgc	acgatataca	ggattttgcc	aaagggttcg	16260
tgtagacttt	ccttggtgta	tccaacggcg	tcagccgggc	aggataggtg	aagtaggccc	16320
acccgcgagc	gggtgttcct	tcttcactgt	cccttattcg	cacctggcgg	tgctcaacgg	16380
gaatcctgct	ctgcgaggct	ggccggctac	cgccggcgta	acagatgagg	gcaagcggat	16440
ggctgatgaa	accaagccaa	ccaggaaggg	cagcccacct	atcaaggtgt	actgccttcc	16500
agacgaacga	agagcgattg	aggaaaaggc	ggeggeggee	ggcatgagcc	tgtcggccta	16560
cctgctggcc	gtcggccagg	gctacaaaat	cacgggcgtc	gtggactatg	agcacgtccg	16620
cgagctggcc	cgcatcaatg	gcgacctggg	ccgcctgggc	ggcctgctga	aactctggct	16680
caccgacgac	ccgcgcacgg	cgcggttcgg	tgatgccacg	atcctcgccc	tgctggcgaa	16740
gatcgaagag	aagcaggacg	agcttggcaa	ggtcatgatg	ggcgtggtcc	gcccgagggc	16800
agagccatga	cttttttagc	cgctaaaacg	gccggggggt	gcgcgtgatt	gccaagcacg	16860
tecceatgeg	ctccatcaag	aagagcgact	tcgcggagct	ggtgaagtac	atcaccgacg	16920
agcaaggcaa	gaccgagcgc	ctttgcgacg	ctca			16954

<210> 44 <211> 16954 <212> DNA <213> Artificial <220> <223> Plasmid <220> <221> misc_feature <222> (10264)..(10264) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10472)..(10472) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10563)..(10563) <223> n is a, c, g, or t <400> 44 ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 120 180 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 240 ccggcgcgc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 300 cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 360 420 gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat

tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt

480

WO 2004/063359 PCT/EP2004/000099 199/357

ttttcggcca ccgctaac	ct gtcttttaac	ctgcttttaa	accaatattt	ataaaccttg	540
tttttaacca gggctgcg	ce etgtgegegt	gaccgcgcac	gccgaagggg	ggtgccccc	600
cttctcgaac cctcccgg	cc cgctaacgcg	ggcctcccat	cccccaggg	gctgcgcccc	660
teggeegega aeggeete	ac cccaaaaatg	gcagcgctgg	cagtccttgc	cattgccggg	720
atcggggcag taacggga	tg ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc tggcatcg	ac attcagcgac	caggtgccgg	gcagtgaggg	cggcggcctg	840
ggtggeggee tgeeette	ac ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaattttta ccttgggc	at tettggeata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa acccagcg	gaa ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg gaccttta	ıca gaattactct	: atgaagcgcc	atatttaaaa	agctaccaag	1080
acgaagagga tgaagagg	gat gaggaggcag	g attgccttga	atatattgac	aatactgata	1140
agataatata tetttta	tat agaagatato	gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg cgcttate	caa tatatctata	a gaatgggcaa	. agcataaaaa	a cttgcatgga	1260
ctaatgcttg aaaccca	gga caataacct	t atagcttgta	aattctatca	a taattgggta	1320
atgactccaa cttattg	ata gtgtttat	g ttcagataat	: gcccgatgac	tttgtcatgc	1380
agctccaccg attttga	gaa cgacagcga	c ttccgtccca	ı gccgtgccaç	g gtgetgeete	1440
agattcaggt tatgccg	ctc aattcgctg	c gtatatcgct	tgctgatta	c gtgcagcttt	1500
cccttcaggc gggattc	ata cageggeca	g ccatccgtca	a tccatatca	c cacgtcaaag	1560
ggtgacagca ggctcat	aag acgccccag	c gtcgccata	g tgegtteac	c gaatacgtgc	1620
gcaacaaccg tettecg	gag actgtcata	c gcgtaaaac	a gccagcgct	g gegegattta	1680
gccccgacat agcccca	actg ttcgtccat	t teegegeag	a cgatgacgt	c actgcccggc	1740

WO 2004/063359 PCT/EP2004/000099 200/357

tgtatgcgcg	aggttaccga	ctgcggcctg	agttttttaa	gtgacgtaaa	atcgtgttga	1800
ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg	agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc	cgtcgatact	atgttatacg	ccaactttga	aaacaacttt	gaaaaagctg	2160
ttttctggta	tttaaggttt	tagaatgcaa	ggaacagtga	attggagttc	gtcttgttat	2220
aattagcttc	ttggggtatc	tttaaatact	gtagaaaaga	ggaaggaaat	aataaatggc	2280
taaaatgaga	atatcaccgg	aattgaaaaa	actgatcgaa	aaataccgct	gcgtaaaaga	2340
tacggaagga	atgtctcctg	ctaaggtata	taagctggtg	ggagaaaatg	aaaacctata	2400
tttaaaaatg	acggacagcc	ggtataaagg	gaccacctat	gatgtggaac	gggaaaagga	2460
catgatgcta	tggctggaag	gaaagctgcc	tgttccaaag	gtcctgcact	ttgaacggca	2520
tgatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac	ttactgaata	. acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt	aaagateege	gcgagctgta	. tgattttta	aagacggaaa	agcccgaaga	2820
ggaacttgto	: ttttcccacg	gegaeetggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc	: tttattgato	: ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc	cggtcgatca	ı gggaggatat	: cggggaagaa	cagtatgtcg	agctattttt	3000

WO 2004/063359 PCT/EP2004/000099 201/357

	gggatcaagc	ataattaaaa	maaaataaaa	tattatattt	tactogatoa	3060
tgacttactg	gggatcaage	CLYALLYYYA	gaaaacaaaa		ouougguugu	
attgttttag	tacctagatg	tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tcttccgcat	caagtgtttt	ggctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt	attcgtgcag	ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg	gaccgacttc	attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa	tcaggaataa	gggcacattg	ccccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc	cgaggatgcc	gaaaccatcg	caagccgcac	cgtcatgcgt	gcgccccgcg	3480
aaaccttcca	gtccgtcggc	tcgatggtcc	agcaagctac	ggccaagatc	gagegegaca	3540
gcgtgcaact	ggctccccct	gccctgcccg	cgccatcggc	cgccgtggag	cgttcgcgtc	3600
gtctcgaaca	ggaggcggca	ggtttggcga	agtcgatgac	catcgacacg	cgaggaacta	3660
tgacgaccaa	gaagcgaaaa	accgccggcg	aggacctggc	aaaacaggtc	agcgaggcca	3720
agcaggccgc	gttgctgaaa	cacacgaagc	agcagatcaa	ggaaatgcag	ctttccttgt	3780
tcgatattgc	gccgtggccg	gacacgatgc	gagcgatgcc	aaacgacacg	gcccgctctg	3840
ccctgttcac	cacgcgcaac	aagaaaatcc	cgcgcgaggc	gctgcaaaac	aaggtcattt	3900
tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctatc	ggcgagccga	4020
tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggo	: cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg	gcacctggaa	. teggtgtege	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200
gcaagaaaac	: gtcccgttgc	caggtcctga	ı tegaegagga	aatcgtcgtg	ctgtttgctg	4260

WO 2004/063359 PCT/EP2004/000099 203/357

agcgggttta	acctacttcc	tttggttccg	ggggateteg	cgactcgaac	ctacagttgt	5580
ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	tgcatcaggc	5640
cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	ataggggagt	5700
tgatatcgtc	aacgttcact	tctaaagaaa	tagegeeact	cagcttcctc	agcggcttta	5760
tccagcgatt	tcctattatg	tcggcatagt	tctcaagatc	gacagcctgt	cacggttaag	5820
cgagaaatga	ataagaaggc	tgataattcg	gatctctgcg	agggagatga	tatttgatca	5880
caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	gatcatccgt	5940
gtttcaaacc	cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	atgagcaaag	6000
tctgccgcct	tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	ctgcctgtat	6060
cgagtggtga	ttttgtgccg	agctgccggt	cggggagctg	ttggctggct	ggtggcagga	6120
tatattgtgg	tgtaaacaaa	ttgacgctta	gacaacttaa	taacacattg	cggacgtttt	6180
taatgtactg	gggtggtttt	tcttttcacc	agtgagacgg	gcaacagctg	attgcccttc	6240
accgcctggc	cctgagagag	ttgcagcaag	cggtccacgc	tggtttgccc	cagcaggcga	6300
aaatcctgtt	tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	6360
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480
cacccaaatc	aagtttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	6540
ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600
agaaagcgaa	aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt	cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720
tgggtaacgc	: cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780

ctcggtaccc	ggggatcttt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840
cgaggagcct	cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900
gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac	gctcttctga	tccagtcgat	catccgctga	aggcgctttc	gaatctggtt	7020
aagatccacg	tcttcgggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140
aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag	aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320
tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
tcgaaataca	teeggatgte	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560
tagcagttco	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	ı gctcaccaca	aaagtcagac	: ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggatti	: cggcacggct	acggaagacg	gagaagcca	c cttcagtgga	a ctcgagtacc	7740
atttaattc	atttgtgttt	. gatcgagaco	taatacagco	c cctacaacga	a ccatcaaagt	7800
cgtatagct	a ccagtgagga	agtggactca	a aatcgactto	agcaacatch	cctggataaa	7860
ctttaagcc	t aaactataca	ı gaataagatı	a ggtggagag	ttataccga	g ctcccaaatc	7920
tgtccagat	c atggttgaco	ggtgcctgg	a tcttcctate	a gaatcatcc	t tattcgttga	7980
cctagctga	t tctggagtga	a cccagaggg	t catgacttg	a gcctaaaat	c cgccgcctcc	8040

WO 2004/063359 PCT/EP2004/000099 205/357

accatttgta	gaaaaatgtg	acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100
tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataagc	cactggccag	8160
acagctctgg	cggctctgag	gtgcagtgga	tgattattaa	teegggaeeg	geegeeete	8220
cgccccgaag	tggaaaggct	ggtgtgcccc	tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatagccgt	cggcgaaata	gcatgccatt	aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	agattcacga	gatagtacct	tctccgaagt	aggtagagcg	agtacccggc	8460
gcgtaagctc	cctaattggc	ccatccggca	tctgtagggc	gtccaaatat	cgtgcctctc	8520
ctgctttgcc	cggtgtatga	aaccggaaag	gccgctcagg	agctggccag	cggcgcagac	8580
cgggaacaca	agctggcagt	cgacccatcc	ggtgctctgc	actcgacctg	ctgaggtccc	8640
tcagtccctg	gtaggcagct	ttgccccgtc	tgtccgcccg	gtgtgtcggc	ggggttgaca	8700
aggtcgttgc	gtcagtccaa	catttgttgc	catattttcc	tgctctcccc	accagctgct	8760
cttttcttt	ctctttcttt	tcccatcttc	agtatattca	tcttcccatc	caagaacctt	8820
tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300

WO 2004/063359 PCT/EP2004/000099 206/357

ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	gcccattcgg	9360
accgcaagga	atcggtcaat	acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	tegegeagge	9480
tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	tegtgeaege	9540
ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	cggagcttgc	9720
aggategeeg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga	gccgggactg	tegggegtac	acaaatcgcc	cgcagaagcg	cggccgtctg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960.
gagggcaaag	gaatagagta	gatgccgacc	gegggatega	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080
ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	ı gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttcgggttta	10200
cctcttccag	g atacagetea	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagago	: tattttcatt	: ttcgggagac	10320
gagatcaago	c agatcaacgg	, tegtcaagag	acctacgaga	a ctgaggaato	c cgctcttggc	10380
tccacgcga	c tatatatttg	, tctctaattg	tactttgace	a tgctcctctt	ctttactctg	10440
atagettga	c tatgaaaatt	: ccgtcaccag	cncetgggtt	: cgcaaagata	a attgcatgtt	10500
tetteettg	a actctcaago	c ctacaggaca	cacattcato	c gtaggtataa	a acctcgaaat	10560

WO 2004/063359 PCT/EP2004/000099 207/357

canttectae taagatggta tacaatagta accatgeatg gttgeetagt gaatgeteeg 10620 taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttacccaga 10680 atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740 gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt agagataaaa 10800 taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac atacacgcta 10860 tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc acaagattat 10920 ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca aagatactgt 10980 catctcgaga gcgttgcaca ggctgctgtt tgccaaattg gatgtttgcc gaattagtaa 11040 aatacgcaag catttettac ettteegete eettteeta atteteecaa agactaaatg 11100 aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga tataaacttg 11160 cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaaacatg tataaaaaaa 11220 aactttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat tgcttgtata 11280 tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta tttataaaaa 11340 gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct attttaatct 11400 catgttgtgg atttggaatg ccctgatcgt tttcgttacc gtgattggca tggaagtgat 11460 tgctgcactg gcacacaaat acatcatgca cggctggggt tggggatggc atctttcaca 11520 tcatgaaccg cgtaaaggtg cgtttgaagt taacgatctt tatgccgtgg tttttgctgc 11580 attatcgatc ctgctgattt atctgggcag tacaggaatg tggccgctcc agtggattgg 11640 cgcaggtatg acggcgtatg gattactcta ttttatggtg cacgacgggc tggtgcatca 11700 acgttggcca ttccgctata ttccacgcaa gggctacctc aaacggttgt atatggcgca 11760 ccgtatgcat cacgccgtca ggggcaaaga aggttgtgtt tcttttggct tcctctatgc 11820

WO 2004/063359 PCT/EP2004/000099 208/357

gccgccctg tcaaaacttc aggcgacgct ccgggaaaga catggcgcta gagcgggcgc 11	1880
tgccagagat gcgcagggcg gggaggatga gcccgcatcc gggaagtaag ggcctgacca 11	1940
gaggcggcca gcagcagcgt taatttttcg ggcgtggtcg ttgactgccg ctgatcccaa 12	2000
agcttggcgt aatcatggtc atagctgttt cctgtgtgaa attgttatcc gctcacaatt 12	2060
ccacacaaca tacgagccgg aagcataaag tgtaaagcct ggggtgccta atgagtgagc 12	2120
taactcacat taattgcgtt gcgctcactg cccgctttcc agtcgggaaa cctgtcgtgc 12	2180
cagctgcatt aatgaatcgg ccaacgcgcg gggagaggcg gtttgcgtat tgggccaaag 12	2240
acaaaagggc gacattcaac cgattgaggg agggaaggta aatattgacg gaaattattc 12	2300
attaaaggtg aattatcacc gtcaccgact tgagccattt gggaattaga gccagcaaaa 12	2360
tcaccagtag caccattacc attagcaagg ccggaaacgt caccaatgaa accatcgata 12	2420
gcagcaccgt aatcagtagc gacagaatca agtttgcctt tagcgtcaga ctgtagcgcg 12	2480
ttttcatcgg cattttcggt catagccccc ttattagcgt ttgccatctt ttcataatca 12	2540
aaatcaccgg aaccagagcc accaceggaa cegeeteeet cagageegee acceteagaa 1	2600
ccgccaccct cagagccacc accetcagag ccgccaccag aaccaccacc agagccgccg 1	.2660
ccagcattga caggaggccc gatctagtaa catagatgac accgcgcgcg ataatttatc 1	.2720
ctagtttgcg cgctatattt tgttttctat cgcgtattaa atgtataatt gcgggactct 1	.2780
aatcataaaa acccatctca taaataacgt catgcattac atgttaatta ttacatgctt 1	.2840
aacgtaattc aacagaaatt atatgataat catcgcaaga ccggcaacag gattcaatct 1	2900
taagaaactt tattgccaaa tgtttgaacg atcggggatc atccgggtct gtggcgggaa 1	L2960
ctccacgaaa atatccgaac gcagcaagat atcgcggtgc atctcggtct tgcctgggca 1	13020
gtegeegeeg aegeegttga tgtggaegee gggeeegate atattgtege teaggategt 1	L3080

WO 2004/063359 PCT/EP2004/000099 209/357

ggcgttgtgc	ttgtcggccg	ttgctgtcgt	aatgatatcg	gcaccttcga	ccgcctgttc	13140
cgcagagatc	ccgtgggcga	agaactccag	catgagatcc	ccgcgctgga	ggatcatcca	13200
gccggcgtcc	cggaaaacga	ttccgaagcc	caacctttca	tagaaggcgg	cggtggaatc	13260
gaaatctcgt	gatggcaggt	tgggcgtegc	ttggtcggtc	atttcgaacc	ccagagtccc	13320
gctcagaaga	actcgtcaag	aaggcgatag	aaggcgatgc	gctgcgaatc	gggagcggcg	13380
ataccgtaaa	gcacgaggaa	gcggtcagcc	cattcgccgc	caagctcttc	agcaatatca	13440
cgggtagcca	acgctatgtc	ctgatagcgg	teegecacae	ccagccggcc	acagtcgatg	13500
aatccagaaa	agcggccatt	ttccaccatg	atattcggca	agcaggcatc	gccatgggtc	13560
acgacgagat	catcgccgtc	gggcatgcgc	gccttgagcc	tggcgaacag	tteggetgge	13620
gcgagcccct	gatgctcttc	gtccagatca	tectgatega	caagaccggc	ttccatccga	13680
gtacgtgctc	gctcgatgcg	atgtttcgct	tggtggtcga	atgggcaggt	agccggatca	13740
agcgtatgca	gccgccgcat	tgcatcagcc	atgatggata	ctttctcggc	aggagcaagg	13800
tgagatgaca	ggagatcctg	ccccggcact	tegeceaata	gcagccagtc	ccttcccgct	13860
tcagtgacaa	cgtcgagcac	agctgcgcaa	ggaacgcccg	tegtggccag	ccacgatagc	13920
cgcgctgcct	cgtcctgcag	ttcattcagg	gcaccggaca	ggtcggtctt	gacaaaaaga	13980
accgggcgcc	cctgcgctga	cagceggaae	acggcggcat	cagagcagcc	gattgtctgt	14040
tgtgcccagt	catagccgaa	tagcetetee	acccaagcgg	ccggagaacc	tgcgtgcaat	14100
ccatcttgtt	caatcatgcg	aaacgatcca	gatccggtgc	agattatttg	gattgagagt	14160
gaatatgaga	ctctaattgg	ataccgaggg	gaatttatgg	aacgtcagtg	gagcattttt	14220
gacaagaaat	atttgctagc	tgatagtgac	cttaggcgac	: ttttgaacgo	gcaataatgg	14280
tttctgacgt	atgtgcttag	ctcattaaac	tecagaaaco	: cgcggctgag	g tggctccttc	14340

	-t-at-at-az-at	tccaaacgta :	a a a concetto	tecegeatea	tcaacaaaaa	14400
aacgttgcgg	ctctgtcagt	cccaaacyta	addeggeoog			
tcataacgtg	actcccttaa	ttctccgctc	atgatcagat	tgtcgtttcc	cgccttcagt	14460
ttaaactatc	agtgtttgac	aggatatatt	ggcgggtaaa	cctaagagaa	aagagcgttt	14520
attagaataa	teggatattt	aaaagggcgt	gaaaaggttt	atccgttcgt	ccatttgtat	14580
gtgcatgcca	accacagggt	tecceagate	tggcgccggc	cagcgagacg	agcaagattg	14640
geegeegeee	gaaacgatcc	gacagcgcgc	ccagcacagg	tgcgcaggca	aattgcacca	14700
acgcatacag	cgccagcaga	atgccatagt	gggcggtgac	gtcgttcgag	tgaaccagat	14760
cgcgcaggag	gcccggcagc	accggcataa	tcaggccgat	gccgacagcg	tcgagcgcga	14820
cagtgctcag	aattacgatc	aggggtatgt	tgggtttcac	gtctggcctc	cggaccagcc	14880
tccgctggtc	cgattgaacg	cgcggattct	ttatcactga	taagttggtg	gacatattat	14940
gtttatcagt	gataaagtgt	caagcatgac	aaagttgcag	ccgaatacag	tgatccgtgc	15000
cgccctggac	ctgttgaacg	aggtcggcgt	agacggtctg	acgacacgca	aactggcgga	15060
acggttgggg	gttcagcagc	cggcgcttta	ctggcacttc	aggaacaagc	gggcgctgct	15120
cgacgcactg	gccgaagcca	tgctggcgga	gaatcatacg	catteggtge	cgagagccga	15180
cgacgactgg	cgctcatttc	tgatcgggaa	tgcccgcagc	ttcaggcagg	cgctgctcgc	15240
ctacegegat	ggcgcgcgca	tccatgccgg	cacgcgaccg	ggcgcaccgc	agatggaaac	15300
ggccgacgcg	g cagetteget	tcctctgcga	ggcgggtttt	teggeegggg	acgccgtcaa	15360
tgcgctgatg	g acaatcagct	acttcactgt	tggggccgtg	cttgaggago	: aggccggcga	15420
cagcgatgco	ggcgagcgcg	gcggcaccgt	tgaacaggct	cegetetege	cgctgttgcg	15480
ggccgcgata	a gacgccttcg	acgaagccgg	tccggacgca	a gegttegage	agggactcgc	15540
ggtgattgt	gatggattgg	g cgaaaaggag	gctcgttgtc	aggaacgtt	g aaggaccgag	15600

WO 2004/063359 PCT/EP2004/000099 211/357

aaagggtgac	gattgatcag	gaccgctgcc	ggagcgcaac	ccactcacta	cagcagagcc	15660
atgtagacaa	catcccctcc	ccctttccac	cgcgtcagac	gcccgtagca	gcccgctacg	15720
ggctttttca	tgccctgccc	tagegtecaa	gcctcacggc	cgcgctcggc	ctctctggcg	15780
gccttctggc	gctcttccgc	ttcctcgctc	actgactcgc	tgcgctcggt	cgttcggctg	15840
cggcgagcgg	tatcagctca	ctcaaaggcg	gtaatacggt	tatccacaga	atcaggggat	15900
aacgcaggaa	agaacatgtg	agcaaaaggc	cagcaaaagg	ccaggaaccg	taaaaaggcc	15960
gcgttgctgg	cgtttttcca	taggctccgc	cccctgacg	agcatcacaa	aaatcgacgc	16020
tcaagtcaga	ggtggcgaaa	cccgacagga	ctataaagat	accaggcgtt	tccccctgga	16080
agctccctcg	tgcgctctcc	tgttccgacc	ctgccgctta	ccggatacct	gteegeettt	16140
ctcccttcgg	gaagcgtggc	gcttttccgc	tgcataaccc	tgcttcgggg	tcattatagc	16200
gattttttcg	gtatatccat	cctttttcgc	acgatataca	ggattttgcc	aaagggttcg	16260
tgtagacttt	ccttggtgta	tccaacggcg	tcagccgggc	aggataggtg	aagtaggccc	16320
acccgcgagc	gggtgttcct	tcttcactgt	cccttattcg	cacctggcgg	tgctcaacgg	16380
gaatcctgct	ctgcgaggct	ggccggctac	cgccggcgta	acagatgagg	gcaagcggat	16440
ggctgatgaa	accaagccaa	ccaggaaggg	cagcccacct	atcaaggtgt	actgccttcc	16500
agacgaacga	agagcgattg	aggaaaaggc	ggcggcggcc	ggcatgagcc	tgtcggccta	16560
cctgctggcc	gtcggccagg	gctacaaaat	cacgggcgtc	gtggactatg	agcacgtccg	16620
cgagetggee	cgcatcaatg	gcgacctggg	ccgcctgggc	ggcctgctga	aactctggct	16680
caccgacgac	ccgcgcacgg	cgcggttcgg	tgatgccacg	atcctcgccc	tgctggcgaa	16740
gatcgaagag	aagcaggacg	agcttggcaa	ggtcatgatg	ggcgtggtcc	gcccgagggc	16800
agagccatga	cttttttage	cgctaaaacg	gccggggggt	gcgcgtgatt	gccaagcacg	16860

tccccatgcg ctccatcaag aagagcgact tcgcggagct ggtgaagtac atcaccgacg 16920

agcaaggcaa gaccgagcgc ctttgcgacg ctca 16954

<210> 45

<211> 19491

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (18970)..(18970)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19178)..(19178)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (19269)..(19269)

<223> n is a, c, g, or t

<400> 45

agettggtac egagetegga tecactagta aeggeegeea gtgtgetgga attegeeett 60

gaeggeeagt gaattegage teggtaceeg gggatettte gaeaetgaaa taegtegage 120

etgeteeget tggaagegge gaggageete gteetgteae aaetaceaae atggagtaeg 180

ataagggeea gtteegeeag etcattaaga geeagtteat gggegttgge atgatggeeg 240

teatgeatet gtaetteaag taeaeeaaeg etettetgat eeagtegate ateegetgaa 300

ggcgctttcg aatctggtta agatccacgt cttcgggaag ccagcgactg gtgacctcca

360

WO 2004/063359 PCT/EP2004/000099 213/357

gcgtcccttt	aaggctgcca	acagctttct	cagccagggc	cagcccaaga	Ccgacaaggc	420
ctccctccag	aacgccgaga	agaactggag	gggtggtgtc	aaggaggagt	aagctcctta	480
ttgaagtcgg	aggacggagc	ggtgtcaaga	ggatattctt	cgactctgta	ttatagataa	540
gatgatgagg	aattggaggt	agcatagett	catttggatt	tgctttccag	gctgagactc	600
tagcttggag	catagagggt	cctttggctt	tcaatattct	caagtatctc	gagtttgaac	660
ttattccctg	tgaacctttt	attcaccaat	gagcattgga	atgaacatga	atctgaggac	720
tgcaatcgcc	atgaggtttt	cgaaatacat	ccggatgtcg	aaggcttggg	gcacctgcgt	780
tggttgaatt	tagaacgtgg	cactattgat	catccgatag	ctctgcaaag	ggcgttgcac	840
aatgcaagtc	aaacgttgct	agcagttcca	ggtggaatgt	tatgatgagc	attgtattaa	900
atcaggagat	atagcatgat	ctctagttag	ctcaccacaa	aagtcagacg	gcgtaaccaa	960
aagtcacaca	acacaagctg	taaggatttc	ggcacggcta	cggaagacgg	agaagccacc	1020
ttcagtggac	tcgagtacca	tttaattcta	tttgtgtttg	atcgagacct	aatacagccc	1080
ctacaacgac	catcaaagtc	gtatagctac	cagtgaggaa	gtggactcaa	atcgacttca	1140
gcaacatctc	ctggataaac	tttaagccta	aactatacag	aataagatag	gtggagagct	1200
tataccgagc	tcccaaatct	gtccagatca	tggttgaccg	gtgcctggat	cttcctatag	1260
aatcatcctt	attcgttgac	ctagctgatt	ctggagtgac	ccagagggtc	atgacttgag	1320
cctaaaatcc	geegeeteea	ccatttgtag	aaaaatgtga	cgaactcgtg	agctctgtac	1380
agtgaccggt	gactctttct	ggcatgcgga	gagacggacg	gacgcagaga	gaagggctga	1440
gtaataagcc	actggccaga	cagctctggc	ggctctgagg	tgcagtggat	gattattaat	1500
ccgggaccgg	ccgcccctcc	gccccgaagt	ggaaaggctg	gtgtgcccct	cgttgaccaa	1560
gaatctattg	catcatcgga	gaatatggag	cttcatcgaa	tcaccggcag	taagcgaagg	1620

WO 2004/063359 PCT/EP2004/000099 214/357

agaatgtgaa gccaggggtg tatagccgtc ggcgaaatag catgccatta acctaggtac	1680
agaagtccaa ttgcttccga tctggtaaaa gattcacgag atagtacctt ctccgaagta	1740
ggtagagega gtacceggeg egtaagetee etaattggee cateeggeat etgtagggeg	1800
tocaaatato gtgoototoo tgotttgooo ggtgtatgaa accggaaagg cogotcagga	1860
gctggccagc ggcgcagacc gggaacacaa gctggcagtc gacccatccg gtgctctgca	1920
ctcgacctgc tgaggtccct cagtccctgg taggcagctt tgccccgtct gtccgcccgg	1980
tgtgtcggcg gggttgacaa ggtcgttgcg tcagtccaac atttgttgcc atattttcct	2040
geteteceea ecagetgete ttttettte tetttettt eccatettea gtatatteat	2100
cttcccatcc aagaaccttt atttccccta agtaagtact ttgctacatc catactccat	2160
ccttcccatc ccttattcct ttgaaccttt cagttcgagc tttcccactt catcgcagct	2220
tgactaacag ctaccccgct tgagcagaca tcaccatgct gtcgaagctg cagtcaatca	2280
gcgtcaaggc ccgccgcgtt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc	2340
atgctcagcg gtgctcgtta gttcggctgc gagtggcagc accacagaca gaggaggcgc	2400
tgggaaccgt gcaggctgcc ggcgcgggcg atgagcacag cgccgatgta gcactccagc	2460
agettgaceg ggetategea gagegtegtg ceeggegeaa aegggageag etgteatace	2520
aggetgeege cattgeagea teaattggeg tgteaggeat tgccatette geeacetace	2580
tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atggggtgaa gtggctggca	2640
ctctcctctt ggtggttggt ggcgcgctcg gcatggagat gtatgcccgc tatgcacaca	2700
aagccatctg gcatgagtcg cctctgggct ggctgctgca caagagccac cacacacctc	2760
gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc	2820
tectgtgtac etttggette tggetgeeca acgteetggg ggeggeetge tttggagegg	2880

WO 2004/063359 PCT/EP2004/000099 215/357

ggctgggcat	cacgctatac	ggcatggcat	atatgtttgt	acacgatggc	ctggtgcaca	2940
ggcgctttcc	cacegggcce	atcgctggcc	tgccctacat	gaagcgcctg	acagtggccc	3000
accagctaca	ccacagcggc	aagtacggtg	gegegeeetg	gggtatgttc	ttgggtccac	3060
aggagctgca	gcacattcca	ggtgcggcgg	aggaggtgga	gcgactggtc	ctggaactgg	3120
actggtccaa	gcggtagggt	gcggaaccag	gcacgctggt	ttcacacctc	atgcctgtga	3180
taaggtgtgg	ctagagcgat	gcgtgtgaga	cgggtatgtc	acggtcgact	ggtctgatgg	3240
ccaatggcat	cggccatgtc	tggtcatcac	gggctggttg	cctgggtgaa	ggtgatgcac	3300
atcatcatgt	gcggttggag	gggctggcac	agtgtgggct	gaactggagc	agttgtccag	3360
gctggcgttg	aatcagtgag	ggtttgtgat	tggcggttgt	gaagcaatga	ctccgcccat	3420
attctatttg	tgggagctga	gatgatggca	tgcttgggat	gtgcatggat	catggtagtg	3480
cagcaaacta	tattcaccta	gggctgttgg	taggatcagg	tgaggccttg	cacattgcat	3540
gatgtactcg	tcatggtgtg	ttggtgagag	gatggatgtg	gatggatgtg	tattctcaga	3600
cgtagacctt	gactggaggc	ttgatcgaga	gagtgggccg	tattctttga	gaggggaggc	3660
tegtgecaga	aatggtgagt	ggatgactgt	gacgctgtac	attgcaggca	ggtgagatgc	3720
actgtctcga	ttgtaaaata	cattcagatg	caagcttggc	gtaatcatgg	tcatagctgt	3780
ttcctgtgtg	aaattgttat	ccgctcacaa	ttccacacaa	catacgagcc	ggaagcataa	3840
agtgtaaagc	ctggggtgcc	taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	3900
tgcccgcttt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	ggccaacgcg	3960
cggggagagg	cggtttgcgt	attgggccaa	agacaaaagg	gcgacattca	accgattgag	4020
ggagggaagg	taaatattga	cggaaattat	tcattaaagg	tgaattatca	ccgtcaccga	4080
cttgagccat	ttgggaatta	gagccagcaa	aatcaccagt	agcaccatta	ccattagcaa	4140

WO 2004/063359 PCT/EP2004/000099 216/357

ggccggaaac	gtcaccaatg	aaaccatcga	tagcagcacc	gtaatcagta	gcgacagaat	4200
caagtttgcc	tttagcgtca	gactgtagcg	cgttttcatc	ggcattttcg	gtcatagccc	4260
ccttattagc	gtttgccatc	ttttcataat	caaaatcacc	ggaaccagag	ccaccaccgg	4320
aaccgcctcc	ctcagagccg	ccaccctcag	aaccgccacc	ctcagagcca	ccaccctcag	4380
agccgccacc	agaaccacca	ccagagccgc	cgccagcatt	gacaggaggc	ccgatctagt	4440
aacatagatg	acaccgcgcg	cgataattta	tcctagtttg	cgcgctatat	tttgttttct	4500
atcgcgtatt	aaatgtataa	ttgcgggact	ctaatcataa	aaacccatct	cataaataac	4560
gtcatgcatt	acatgttaat	tattacatgc	ttaacgtaat	tcaacagaaa	ttatatgata	4620
atcatcgcaa	gaccggcaac	aggattcaat	cttaagaaac	tttattgcca	aatgtttgaa	4680
cgatcgggga	tcatccgggt	ctgtggcggg	aactccacga	aaatatccga	acgcagcaag	4740
atatcgcggt	gcatctcggt	cttgcctggg	cagtcgccgc	cgacgccgtt	gatgtggacg	4800
ccgggcçcga	tcatattgtc	gctcaggatc	gtggcgttgt	gcttgtcggc	cgttgctgtc	4860
[*] gtaatgatat	cggcaccttc	gaccgcctgt	tccgcagaga	tecegtggge	gaagaactcc	4920
agcatgagat	ccccgcgctg	gaggatcatc	cagccggcgt	cccggaaaac	gattccgaag	4980
cccaaccttt	catagaaggc	ggcggtggaa	tcgaaatctc	gtgatggcag	gttgggcgtc	5040
gcttggtcgg	tcatttcgaa	ccccagagtc	ccgctcagaa	gaactcgtca	agaaggcgat	5100
agaaggcgat	gegetgegaa	tcgggagcgg	cgataccgta	. aagcacgagg	g aageggteag	5160
cccattcgcc	gccaagctct	tcagcaatat	cacgggtago	caacgctatg	g teetgatage	5220
ggtccgccac	: acccagccgg	ccacagtcga	tgaatccaga	aaagcggcca	a ttttccacca	5280
tgatattcgg	g caagcaggca	tegecatggg	tcacgacgag	atcatcgccg	g tegggeatge	5340
gcgccttgag	g cctggcgaac	agttcggctg	gegegageed	ctgatgctct	tegtecagat	5400

WO 2004/063359 PCT/EP2004/000099 217/357

catcctgatc	gacaagaccg	gcttccatcc	gagtacgtgc	tegetegatg	cgatgtttcg	5460
cttggtggtc	gaatgggcag	gtagccggat	caagcgtatg	cageegeege	attgcatcag	5520
ccatgatgga	tactttctcg	gcaggagcaa	ggtgagatga	caggagatcc	tgccccggca	5580
cttcgcccaa	tagcagccag	tecetteceg	cttcagtgac	aacgtcgagc	acagetgege	5640
aaggaacgcc	cgtcgtggcc	agccacgata	gccgcgctgc	ctcgtcctgc	agttcattca	5700
gggcaccgga	caggtcggtc	ttgacaaaaa	gaaccgggcg	ccctgcgct	gacagccgga	5760
acacggcggc	atcagagcag	ccgattgtct	gttgtgccca	gtcatagccg	aatagcctct	5820
ccacccaagc	ggccggagaa	cctgcgtgca	atccatcttg	ttcaatcatg	cgaaacgatc	5880
cagatccggt	gcagattatt	tggattgaga	gtgaatatga	gactctaatt	ggataccgag	5940
gggaatttat	ggaacgtcag	tggagcattt	ttgacaagaa	atatttgcta	gctgatagtg	6000
accttaggcg	acttttgaac	gcgcaataat	ggtttctgac	gtatgtgctt	agctcattaa	6060
actccagaaa	cccgcggctg	agtggctcct	tcaacgttgc	ggttctgtca	gttccaaacg	6120
taaaacggct	tgtcccgcgt	catcggcggg	ggtcataacg	tgactccctt	aattctccgc	6180
tcatgatcag	attgtcgttt	cccgccttca	gtttaaacta	tcagtgtttg	acaggatata	6240
ttggcgggta	aacctaagag	aaaagagcgt	ttattagaat	aatcggatat	ttaaaagggc	6300
gtgaaaaggt	ttatccgttc	gtccatttgt	atgtgcatgc	caaccacagg	gttccccaga	6360
tctggcgccg	gccagcgaga	cgagcaagat	tggccgccgc	ccgaaacgat	ccgacagcgc	6420
gcccagcaca	ggtgcgcagg	caaattgcac	caacgcatac	agcgccagca	gaatgccata	6480
gtgggcggtg	acgtcgttcg	agtgaaccag	ategegeagg	aggcccggca	gcaccggcat	6540
aatcaggccg	atgccgacag	cgtcgagcgc	gacagtgctc	agaattacga	tcaggggtat	6600
gttgggtttc	acgtctggcc	teeggaceag	cctccgctgg	tccgattgaa	cgcgcggatt	6660

ctttatcact	gataagttgg	tggacatatt	atgtttatca	gtgataaagt	gtcaagcatg	6720
acaaagttgc	agccgaatac	agtgatccgt	gccgccctgg	acctgttgaa	cgaggtcggc	6780
gtagacggtc	tgacgacacg	caaactggcg	gaacggttgg	gggttcagca	gccggcgctt	6840
tactggcact	tcaggaacaa	gcgggcgctg	ctcgacgcac	tggccgaagc	catgctggcg	6900
gagaatcata	cgcattcggt	gccgagagcc	gacgacgact	ggcgctcatt	tctgatcggg	6960
aatgcccgca	gcttcaggca	ggcgctgctc	gcctaccgcg	atggcgcgcg	catccatgcc	7020
ggcacgcgac	cgggcgcacc	gcagatggaa	acggccgacg	cgcagcttcg	cttcctctgc	7080
gaggcgggtt	tttcggccgg	ggacgccgtc	aatgegetga	tgacaatcag	ctacttcact	7140
gttggggccg	tgcttgagga	gcaggccggc	gacagċgatg	ccggcgagcg	cggcggcacc	7200
gttgaacagg	ctccgctctc	geegetgttg	cgggccgcga	tagacgcctt	cgacgaagcc	7260
ggtccggacg	cagcgttcga	gcagggactc	gcggtgattg	tcgatggatt	ggcgaaaagg	7320
aggctcgttg	tcaggaacgt	tgaaggaccg	agaaagggtg	acgattgatc	aggaccgctg	7380
ccggagcgca	acccactcac	tacagcagag	ccatgtagac	aacatcccct	cccctttcc	7440
accgcgtcag	acgcccgtag	cagecegeta	cgggcttttt	catgeeetge	cctagcgtcc	7500
aagcctcacg	geegegeteg	gcctctctgg	gecttetg	gegetettee	getteetege	7560
tcactgacto	getgegeteg	gtcgttcggc	: tgcggcgagc	ggtatcagct	cactcaaagg	7620
cggtaatac	g gttatccace	ı gaatcagggg	g ataacgcagg	g aaagaacato	g tgagcaaaag	7680
gccagcaaa	a ggccaggaac	cgtaaaaagg	g ccgcgttgct	ggcgttttt	c cataggetee	7740
gccccctg	a cgagcatcad	c aaaaatcga	gctcaagtca	a gaggtggcga	a aacccgacag	7800
gactataaa	g ataccaggc	g tttccccct	g gaagctccc	t cgtgcgctc	t cctgttccga	7860
ccctgccgc	t taccggata	c ctgtccgcc	t ttctccctt	c gggaagcgt	g gcgcttttcc	7920

WO 2004/063359 PCT/EP2004/000099 219/357

gctgcataac	cctgcttcgg	ggtcattata	gcgattttt	cggtatatcc	atcctttttc	7980
gcacgatata	caggattttg	ccaaagggtt	cgtgtagact	ttccttggtg	tatccaacgg	8040
cgtcagccgg	gcaggatagg	tgaagtaggc	ccacccgcga	gcgggtgttc	cttcttcact	8100
gtcccttatt	cgcacctggc	ggtgctcaac	gggaatcctg	ctctgcgagg	ctggccggct	8160
accgccggcg	taacagatga	gggcaagcgg	atggctgatg	aaaccaagcc	aaccaggaag	8220
ggcagcccac	ctatcaaggt	gtactgcctt	ccagacgaac	gaagagcgat	tgaggaaaag	8280
gcggcggcgg	ccggcatgag	cctgtcggcc	tacctgctgg	ccgtcggcca	gggctacaaa	8340
atcacgggcg	tegtggaeta	tgagcacgtc	cgcgagctgg	cccgcatcaa	tggcgacctg	8400
ggccgcctgg	geggeetget	gaaactctgg	ctcaccgacg	acccgcgcac	ggcgcggttc	8460
ggtgatgcca	cgatcctcgc	cctgctggcg	aagatcgaag	agaagcagga	cgagcttggc	8520
aaggtcatga	tgggcgtggt	ccgcccgagg	gcagagccat	gacttttta	gccgctaaaa	8580
cggccggggg	gtgcgcgtga	ttgccaagca	cgtccccatg	cgctccatca	agaagagcga	8640
cttcgcggag	ctggtgaagt	acatcaccga	cgagcaaggc	aagaccgagc	gcctttgcga	8700
cgctcaccgg	gctggttgcc	ctegeegetg	ggctggcggc	cgtctatggc	cctgcaaacg	8760
cgccagaaac	gccgtcgaag	ccgtgtgcga	gacaccgcgg	ccgccggcgt	tgtggatacc	8820
tcgcggaaaa	cttggccctc	actgacagat	gaggggcgga	cgttgacact	tgaggggccg	8880
actcacccgg	cgcggcgttg	acagatgagg	ggcaggctcg	atttcggccg	gcgacgtgga	8940
gctggccagc	ctcgcaaatc	ggcgaaaacg	cctgatttta	cgcgagtttc	ccacagatga	9000
tgtggacaag	cctggggata	agtgccctgc	ggtattgaca	cttgaggggc	gcgactactg	9060
acagatgagg	ggegegatee	ttgacacttg	aggggcagag	tgctgacaga	tgaggggcgc	9120
acctattgac	atttgagggg	ctgtccacag	gcagaaaatc	cagcatttgc	aagggtttcc	9180

WO 2004/063359 PCT/EP2004/000099 220/357

gcccgttttt	cggccaccgc	taacctgtct	tttaacctgc	ttttaaacca	atatttataa	9240
accttgtttt	taaccagggc	tgegeeetgt	gcgcgtgacc	gcgcacgccg	aaggggggtg	9300
ccccccttc	tcgaaccctc	ccggcccgct	aacgcgggcc	teccatecce	ccaggggctg	9360
cgcccctcgg	ccgcgaacgg	cctcacccca	aaaatggcag	cgctggcagt	ccttgccatt	9420
gccgggatcg	gggcagtaac	gggatgggcg	atcagcccga	gegegaegee	cggaagcatt	9480
gacgtgccgc	aggtgctggc	atcgacattc	agegaceagg	tgccgggcag	tgagggcggc	9540
ggcctgggtg	geggeetgee	cttcacttcg	gccgtcgggg	cattcacgga	cttcatggcg	9600
gggccggcaa	tttttacctt	gggcattctt	ggcatagtgg	tegegggtge	cgtgctcgtg	9660
ttcgggggtg	cgataaaccc	agcgaaccat	ttgaggtgat	aggtaagatt	ataccgaggt	9720
atgaaaacga	gaattggacc	tttacagaat	tactctatga	agcgccatat	ttaaaaagct	9780
accaagacga	agaggatgaa	gaggatgagg	aggcagattg	ccttgaatat	attgacaata	9840
ctgataagat	aatatatctt	ttatatagaa	gatatcgccg	tatgtaagga	tttcaggggg	9900
caaggcatag	gcagcgcgct	tatcaatata	tctatagaat	gggcaaagca	taaaaacttg	9960
catggactaa	tgcttgaaac	ccaggacaat	aaccttatag	cttgtaaatt	ctatcataat	10020
tgggtaatga	ctccaactta	ttgatagtgt	tttatgttca	gataatgccc	gatgactttg	10080
tcatgcagct	ccaccgattt	tgagaacgac	agegaettee	gteceageeg	tgccaggtgc	10140
tgcctcagat	tcaggttatg	ccgctcaatt	cgctgcgtat	atcgcttgct	: gattacgtgc	10200
agctttccct	tcaggcggga	ttcatacago	ggccagccat	: ccgtcatcca	a tatcaccacg	10260
tcaaagggt	g acagcagget	cataagacgo	cccagcgtcg	g ccatagtgcg	g ttcaccgaat	10320
acgtgcgca	a caaccgtctt	ceggagaetg	tcatacgcgt	: aaaacagcca	a gegetggege	10380
gatttagcc	c cgacatagco	: ccactgttcg	tccatttccg	g cgcagacga	t gacgtcactg	10440

WO 2004/063359 PCT/EP2004/000099 221/357

cccggctgta tgcgcgaggt taccgactgc ggcctgagtt ttttaagtga cgtaaaatcg 10500 tgttgaggec aaegeccata atgegggetg ttgeeeggea tecaaegeca tteatggeca 10560 tatcaatgat tttctggtgc gtaccgggtt gagaagcggt gtaagtgaac tgcagttgcc 10620 atgttttacg gcagtgagag cagagatagc gctgatgtcc ggcggtgctt ttgccgttac 10680 gcaccacccc gtcagtagct gaacaggagg gacagctgat agacacagaa gccactggag 10740 cacctcaaaa acaccatcat acactaaatc agtaagttgg cagcatcacc cataattgtg 10800 gtttcaaaat cggctccgtc gatactatgt tatacgccaa ctttgaaaac aactttgaaa 10860 aagctgtttt ctggtattta aggttttaga atgcaaggaa cagtgaattg gagttcgtct 10920 tgttataatt agcttcttgg ggtatcttta aatactgtag aaaagaggaa ggaaataata 10980 aatggctaaa atgagaatat caccggaatt gaaaaaactg atcgaaaaat accgctgcgt 11040 aaaagatacg gaaggaatgt ctcctgctaa ggtatataag ctggtgggag aaaatgaaaa 11100 cctatattta aaaatgacgg acagccggta taaagggacc acctatgatg tggaacggga 11160 aaaggacatg atgctatggc tggaaggaaa gctgcctgtt ccaaaggtcc tgcactttga 11220 acggcatgat ggctggagca atctgctcat gagtgaggcc gatggcgtcc tttgctcgga 11280 agagtatgaa gatgaacaaa gccctgaaaa gattatcgag ctgtatgcgg agtgcatcag 11340 gctctttcac tccatcgaca tatcggattg tccctatacg aatagcttag acagccgctt 11400 agccgaattg gattacttac tgaataacga tctggccgat gtggattgcg aaaactggga 11460 agaagacact ccatttaaag atccgcgcga gctgtatgat tttttaaaga cggaaaagcc 11520 cgaagaggaa cttgtctttt cccacggcga cctgggagac agcaacatct ttgtgaaaga 11580 tggcaaagta agtggcttta ttgatcttgg gagaagcggc agggcggaca agtggtatga 11640 cattgeette tgegteeggt egateaggga ggatateggg gaagaacagt atgtegaget 11700

WO 2004/063359 PCT/EP2004/000099 222/357

attttttgac	ttactgggga	tcaagcctga	ttgggagaaa	ataaaatatt	atattttact	11760
ggatgaattg	ttttagtacc	tagatgtggc	gcaacgatgc	cggcgacaag	caggagcgca	11820
ccgacttctt	ccgcatcaag	tgttttggct	ctcagg ccg a	ggcccacggc	aagtatttgg	11880
gcaaggggtc	gctggtattc	gtgcagggca	agattcggaa	taccaagtac	gagaaggacg	11940
gccagacggt	ctacgggacc	gacttcattg	ccgataaggt	ggattatctg	gacaccaagg	12000
caccaggcgg	gtcaaatcag	gaataagggc	acattgcccc	ggcgtgagtc	ggggcaatcc	12060
cgcaaggagg	gtgaatgaat	cggacgtttg	accggaaggc	atacaggcaa	gaactgatcg	12120
acgcggggtt	ttccgccgag	gatgccgaaa	ccatcgcaag	ccgcaccgtc	atgcgtgcgc	12180
cccgcgaaac	cttccagtcc	gtcggctcga	tggtccagca	agctacggcc	aagatcgagc	12240
gcgacagcgt	gcaactggct	cccctgccc	tgcccgcgcc	atcggccgcc	gtggagcgtt	12300
cgcgtcgtct	cgaacaggag	gcggcaggtt	tggcgaagtc	gatgaccatc	gacacgcgag	12360
gaactatgac	gaccaagaag	cgaaaaaccg	ccggcgagga	cctggcaaaa	caggtcagcg	12420
aggccaagca	ggccgcgttg	ctgaaacaca	cgaagcagca	gatcaaggaa	atgcagcttt	12480
ccttgttcga	tattgcgccg	tggccggaca	cgatgcgagc	gatgccaaac	gacacggccc	12540
gctctgccct	gttcaccacg	cgcaacaaga	aaatcccgcg	cgaggcgctg	caaaacaagg	12600
tcattttcca	cgtcaacaag	gacgtgaaga	tcacctacac	cggcgtcgag	ctgcgggccg	12660
acgatgacga	actggtgtgg	cagcaggtgt	tggagtacgc	gaagcgcacc	cctatcggcg	12720
agccgatcac	cttcacgttc	tacgagcttt	gccaggacct	gggctggtcg	atcaatggcc	12780
ggtattacac	gaaggccgag	gaatgcctgt	cgcgcctaca	ggcgacggcg	atgggcttca	12840
cgtccgaccg	gettgggeae	ctggaatcgg	tgtcgctgct	gcaccgcttc	cgcgtcctgg	12900
accgtggcaa	ı gaaaacgtcc	cgttgccagg	tcctgatcga	cgaggaaatc	gtcgtgctgt	12960

WO 2004/063359 PCT/EP2004/000099 223/357

ttgctggcga	ccactacacg	aaattcatat	gggagaagta	ccgcaagctg	tcgccgacgg	13020
cccgacggat	gttcgactat	ttcagctcgc	accgggagcc	gtacccgctc	aagctggaaa	13080
ccttccgcct	catgtgcgga	tcggattcca	cccgcgtgaa	gaagtggcgc	gagcaggtcg	13140
gcgaagcctg	cgaagagttg	cgaggcagcg	gcctggtgga	acacgcctgg	gtcaatgatg	13200
acctggtgca	ttgcaaacgc	tagggccttg	tggggtcagt	teeggetggg	ggttcagcag	13260
ccagcgcttt	actggcattt	caggaacaag	cgggcactgc	tcgacgcact	tgcttcgctc	13320
agtatcgctc	gggacgcacg	gcgcgctcta	cgaactgccg	ataaacagag	gattaaaatt	13380
gacaattgtg	attaaggctc	agattcgacg	gcttggagcg	gccgacgtgc	aggatttccg	13440
cgagatccga	ttgtcggccc	tgaagaaagc	tccagagatg	ttcgggtccg	tttacgagca	13500
cgaggagaaa	aagcccatgg	aggcgttcgc	tgaacggttg	cgagatgccg	tggcattcgg	13560
cgcctacatc	gacggcgaga	tcattgggct	gtcggtcttc	aaacaggagg	acggccccaa	13620
ggacgctcac	aaggcgcatc	tgtccggcgt	tttcgtggag	cccgaacagc	gaggccgagg	13680
ggtcgccggt	atgctgctgc	gggcgttgcc	ggcgggttta	ttgctcgtga	tgatcgtccg	13740
acagattcca	acgggaatct	ggtggatgcg	catcttcatc	ctcggcgcac	ttaatatttc	13800
gctattctgg	agcttgttgt	ttatttcggt	ctaccgcctg	ccgggcgggg	tegeggegae	13860
ggtaggcgct	gtgcagccgc	tgatggtcgt	gttcatctct	gccgctctgc	taggtagccc	13920
gatacgattg	atggcggtcc	tgggggctat	ttgcggaact	gcgggcgtgg	cgctgttggt	13980
gttgacacca	aacgcagcgc	tagatectgt	cggcgtcgca	gcgggcctgg	cgggggcggt	14040
ttccatggcg	tteggaaceg	tgctgacccg	caagtggcaa	cctcccgtgc	ctctgctcac	14100
ctttaccgcc	tggcaactgg	cggccggagg	acttctgctc	gttccagtag	ctttagtgtt	14160
tgatccgcca	atcccgatgc	ctacaggaac	caatgttctc	ggcctggcgt	ggctcggcct	14220

WO 2004/063359 PCT/EP2004/000099 224/357

gatcggagcg	ggtttaacct	acttcctttg	gttccggggg	atctcgcgac	tcgaacctac	14280
agttgtttcc	ttactgggct	ttctcagccc	cagatctggg	gtcgatcagc	cggggatgca	14340
tcaggccgac	agtcggaact	tegggteece	gacctgtacc	attcggtgag	caatggatag	14400
gggagttgat	atcgtcaacg	ttcacttcta	aagaaatagc	gccactcagc	ttcctcagcg	14460
gctttatcca	gcgatttcct	attatgtcgg	catagttctc	aagatcgaca	gcctgtcacg	14520
gttaagcgag	aaatgaataa	gaaggctgat	aattcggatc	tctgcgaggg	agatgatatt	14580
tgatcacagg	cagcaacgct	ctgtcatcgt	tacaatcaac	atgctaccct	ccgcgagatc	14640
atccgtgttt	caaacccggc	agcttagttg	cegttettec	gaatagcatc	ggtaacatga	14700
gcaaagtctg	ccgccttaca	acggctctcc	cgctgacgcc	gtcccggact	gatgggctgc	14760
ctgtatcgag	tggtgatttt	gtgccgagct	gccggtcggg	gagctgttgg	ctggctggtg	14820
gcaggatata	ttgtggtgta	aacaaattga	cgcttagaca	acttaataac	acattgcgga	14880
cgtttttaat	gtactggggt	ggtttttctt	ttcaccagtg	agacgggcaa	cagctgattg	14940
cccttcaccg	cctggccctg	agagagttgc	agcaagcggt	ccacgctggt	ttgccccagc	15000
aggcgaaaat	cctgtttgat	ggtggttccg	aaatcggcaa	aatcccttat	aaatcaaaag	15060
aatagcccga	. gatagggttg	agtgttgtto	: cagtttggaa	ı caagagtcca	ctattaaaga	15120
acgtggactc	caacgtcaaa	gggcgaaaaa	ccgtctatca	gggcgatggc	ccactacgtg:	15180
aaccatcacc	: caaatcaagt	: tttttggggt	: cgaggtgccg	g taaagcacta	aatcggaacc	15240
ctaaagggag	g cccccgattt	agagettgae	ggggaaagco	ggcgaacgtg	gcgagaaagg	15300
aagggaagaa	a agcgaaagga	a gegggegeea	a ttcaggctgo	gcaactgttg	g ggaagggcga	15360
teggtgeggg	g cetetteget	t attacgccag	g ctggcgaaag	g ggggatgtg	c tgcaaggcga	15420
ttaagttgg	g taacgccag	g gttttccca	g tcacgacgt	t gtaaaacga	ggccagtgaa	15480

WO 2004/063359 PCT/EP2004/000099 225/357

ttegageteg gtaccegggg atetttegae actgaaatae gtegageetg eteegettgg 15540 aagcggcgag gagcctcgtc ctgtcacaac taccaacatg gagtacgata agggccagtt 15600 ccgccagctc attaagagcc agttcatggg cgttggcatg atggccgtca tgcatctgta 15660 cttcaagtac accaacgete ttetgateca gtegateate egetgaagge getttegaat 15720 ctggttaaga tccacgtctt cgggaagcca gcgactggtg acctccagcg tccctttaag 15780 gctgccaaca gctttctcag ccagggccag cccaagaccg acaaggcctc cctccagaac 15840 gccgagaaga actggagggg tggtgtcaag gaggagtaag ctccttattg aagtcggagg 15900 acggagcggt gtcaagagga tattcttcga ctctgtatta tagataagat gatgaggaat 15960 tggaggtagc atagcttcat ttggatttgc tttccaggct gagactctag cttggagcat 16020 agagggtcct ttggctttca atattctcaa gtatctcgag tttgaactta ttccctgtga 16080 accttttatt caccaatgag cattggaatg aacatgaatc tgaggactgc aatcgccatg aggttttcga aatacatccg gatgtcgaag gcttggggca cctgcgttgg ttgaatttag 16200 aacgtggcac tattgatcat ccgatagctc tgcaaagggc gttgcacaat gcaagtcaaa 16260 cgttgctagc agttccaggt ggaatgttat gatgagcatt gtattaaatc aggagatata 16320 gcatgatctc tagttagctc accacaaaag tcagacggcg taaccaaaag tcacacaaca 16380 caagetgtaa ggatttegge aeggetaegg aagaeggaga ageeaeette agtggaeteg 16440 agtaccattt aattotattt gtgtttgatc gagacctaat acagccccta caacgaccat 16500 caaagtcgta tagctaccag tgaggaagtg gactcaaatc gacttcagca acatctcctg 16560 gataaacttt aagcctaaac tatacagaat aagataggtg gagagcttat accgagctcc caaatctgtc cagatcatgg ttgaccggtg cctggatctt cctatagaat catccttatt 16680 cgttgaccta gctgattctg gagtgaccca gagggtcatg acttgagcct aaaatccgcc 16740

WO 2004/063359 PCT/EP2004/000099 226/357

gcctccacca	tttgtagaaa	aatgtgacga	actcgtgagc	tctgtacagt	gaccggtgac	16800
tetttetgge	atgcggagag	acggacggac	gcagagagaa	gggctgagta	ataagccact	16860
ggccagacag	ctctggcggc	tctgaggtgc	agtggatgat	tattaatccg	ggaccggccg	16920
cccctccgcc	ccgaagtgga	aaggctggtg	tgcccctcgt	tgaccaagaa	tctattgcat	16980
catcggagaa	tatggagctt	catcgaatca	ccggcagtaa	gcgaaggaga	atgtgaagcc	17040
aggggtgtat	agccgtcggc	gaaatagcat	gccattaacc	taggtacaga	agtccaattg	17100
cttccgatct	ggtaaaagat	tcacgagata	gtaccttctc	cgaagtaggt	agagcgagta	17160
cccggcgcgt	aagctcccta	attggcccat	ccggcatctg	tagggcgtcc	aaatatcgtg	17220
cctctcctgc	tttgcccggt	gtatgaaacc	ggaaaggccg	ctcaggagct	ggccagcggc	17280
gcagaccggg	aacacaagct	ggcagtcgac	ccatccggtg	ctctgcactc	gacctgctga	17340
ggteceteag	tccctggtag	gcagctttgc	cccgtctgtc	cgcccggtgt	gtcggcgggg	17400
ttgacaaggt	cgttgcgtca	gtccaacatt	tgttgccata	ttttcctgct	ctccccacca	17460
gctgctcttt	tettttetet	ttcttttccc	atcttcagta	tattcatctt	cccatccaag	17520
aacctttatt	tecectaagt	aagtactttg	ctacatccat	actccatcct	tcccatccct	17580
tattcctttg	aacctttcag	ttcgagcttt	cccacttcat	cgcagcttga	ctaacagcta	17640
ccccgcttga	gcagacatca	ccatgcctga	actcaccgcg	acgtctgtcg	agaagtttct	17700
gatcgaaaag	ttcgacagcg	tctccgacct	gatgcagcto	: tcggagggcg	aagaatctcg	17760
tgctttcagc	ttcgatgtag	gagggcgtgg	atatgtcctg	g cgggtaaata	getgegeega	17820
tggtttctac	aaagatcgtt	: atgtttatcg	gcactttgca	teggeegege	tcccgattcc	17880
ggaagtgctt	gacattgggg	g aattcagcga	a gagcctgacc	tattgcatct	ccegcegtge	17940
acagggtgtc	: acgttgcaag	g acctgcctga	a aaccgaacts	g cccgctgttc	tgcagccggt	18000

WO 2004/063359 PCT/EP2004/000099 227/357

cgcggaggcc	atggatgcga	tegetgegge	cgatcttagc	cagacgagcg	ggttcggccc	18060
attcggaccg	caaggaatcg	gtcaatacac	tacatggcgt	gatttcatat	gcgcgattgc	18120
tgatccccat	gtgtatcact	ggcaaactgt	gatggacgac	accgtcagtg	cgtccgtcgc	18180
gcaggctctc	gatgagctga	tgctttgggc	cgaggactgc	cccgaagtcc	ggcacctcgt	18240
gcacgcggat	ttcggctcca	acaatgtcct	gacggacaat	ggccgcataa	cagcggtcat	18300
tgactggagc	gaggcgatgt	tcggggattc	ccaatacgag	gtcgccaaca	tcttcttctg	18360
gaggccgtgg	ttggcttgta	tggagcagca	gacgcgctac	ttcgagcgga	ggcatccgga	18420
gcttgcagga	tegeegegge	teegggegta	tatgctccgc	attggtcttg	accaactcta	18480
tcagagcttg	gttgacggca	atttcgatga	tgcagcttgg	gcgcagggtc	gatgcgacgc	18540
aatcgtccga	teeggageeg	ggactgtcgg	gcgtacacaa	atcgcccgca	gaagcgcggc	18600
cgtctggacc	gatggctgtg	tagaagtact	cgccgatagt	ggaaaccgac	gccccagcac	18660
tcgtccgagg	gcaaaggaat	agagtagatg	ccgaccgcgg	gatcgatcca	cttaacgtta	18720
ctgaaatcat	caaacagctt	gacgaatctg	gatataagat	cgttggtgtc	gatgtcagct	18780
ccggagttga	gacaaatggt	gttcaggatc	tcgataagat	acgttcattt	gtccaagcag	18840
caaagagtgc	cttctagtga	tttaatagct	ccatgtcaac	aagaataaaa	cgcgttttcg	18900
ggtttacctc	ttccagatac	agctcatctg	caatgcatta	atgcattgac	tgcaacctag	18960
taacgccttn	caggctccgg	cgaagagaag	aatagcttag	cagagctatt	ttcattttcg	19020
ggagacgaga	tcaagcagat	caacggtcgt	caagagacct	acgagactga	ggaatccgct	19080
cttggctcca	cgcgactata	tatttgtctc	taattgtact	ttgacatgct	cctcttcttt	19140
actctgatag	cttgactatg	aaaattccgt	caccagence	tgggttcgca	aagataattg	19200
catgtttctt	ccttgaactc	tcaagcctac	aggacacaca	ttcatcgtag	gtataaacct	19260

cgaaatcant teetactaag atggtataca atagtaacca tgeatggttg cetagtgaat 19320
geteegtaac acceaatacg eeggeegaaa etttttaca acteteetat gagtegttta 19380
eecagaatge acaggtacae ttgtttagag gtaateette tttetageta gaagteeteg 19440
tgtactgtgt aagegeecae teeacatete eactegacet geaggeatge a 19491

<210> 46

<211> 21300

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 46

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60

cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120

cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180

cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240

WO 2004/063359 PCT/EP2004/000099 229/357

tcgggaagcc	agcgactggt	gacctccagc	gtccctttaa	ggctgccaac	agctttctca	300
gccagggcca	gcccaagacc	gacaaggcct	ccctccagaa	cgccgagaag	aactggaggg	360
gtggtgtcaa	ggaggagtaa	gctccttatt	gaagtcggag	gacggagcgg	tgtcaagagg	420
atattcttcg	actctgtatt	atagataaga	tgatgaggaa	ttggaggtag	catagcttca	480
tttggatttg	ctttccaggc	tgagactcta	gcttggagca	tagagggtcc	tttggctttc	540
aatattctca	agtatctcga	gtttgaactt	attccctgtg	aaccttttat	tcaccaatga	600
gcattggaat	gaacatgaat	ctgaggactg	caatcgccat	gaggttttcg	aaatacatcc	660
ggatgtcgaa	ggcttggggc	acctgcgttg	gttgaattta	gaacgtggca	ctattgatca	720
tccgatagct	ctgcaaaggg	cgttgcacaa	tgcaagtcaa	acgttgctag	cagttccagg	780
tggaatgtta	tgatgagcat	tgtattaaat	caggagatat	agcatgatct	ctagttagct	840
caccacaaaa	gtcagacggc	gtaaccaaaa	gtcacacaac	acaagctgta	aggatttcgg	900
cacggctacg	gaagacggag	aagccacctt	cagtggactc	gagtaccatt	taattctatt	960
tgtgtttgat	cgagacctaa	tacagcccct	acaacgacca	tcaaagtcgt	atagctacca	1020
gtgaggaagt	ggactcaaat	cgacttcagc	aacatctcct	ggataaactt	taagcctaaa	1080
ctatacagaa	taagataggt	ggagagctta	taccgagctc	ccaaatctgt	ccagatcatg	1140
gttgaccggt	gcctggatct	tcctatagaa	tcatccttat	tcgttgacct	agctgattct	1200
ggagtgaccc	agagggtcat	gacttgagcc	taaaatccgc	cgcctccacc	atttgtagaa	1260
aaatgtgacg	aactcgtgag	ctctgtacag	tgaccggtga	ctctttctgg	catgcggaga	1320
gacggacgga	cgcagagaga	agggctgagt	aataagccac	tggccagaca	gctctggcgg	1380
ctctgaggtg	cagtggatga	ttattaatcc	gggaccggcc	gccctccgc	cccgaagtgg	1440
aaaggctggt	gtgcccctcg	ttgaccaaga	atctattgca	tcatcggaga	atatggagct	1500

						•
tcatcgaatc	accggcagta	agcgaaggag	aatgtgaagc	caggggtgta	tageegtegg	1560
cgaaatagca	tgccattaac	ctaggtacag	aagtccaatt	gcttccgatc	tggtaaaaga	1620
ttcacgagat	agtaccttct	ccgaagtagg	tagagcgagt	acccggcgcg	taagctccct	1680
aattggccca	tccggcatct	gtagggcgtc	caaatatcgt	gcctctcctg	ctttgcccgg	1740
tgtatgaaac	cggaaaggcc	gctcaggagc	tggccagcgg	cgcagaccgg	gaacacaagc	1800
tggcagtcga	cccatccggt	gctctgcact	cgacctgctg	aggtccctca	gtccctggta	1860
ggcagctttg	ccccgtctgt	ccgcccggtg	tgtcggcggg	gttgacaagg	tegttgegte	1920
agtccaacat	ttgttgccat	attttcctgc	tctccccacc	agctgctctt	ttcttttctc	1980
tttcttttcc	catcttcagt	atattcatct	tcccatccaa	gaacctttat	ttcccctaag	2040
taagtacttt	gctacatcca	tactccatcc	ttcccatccc	ttattccttt	gaacctttca	2100
gttcgagctt	tcccacttca	tegeagettg	actaacagct	accccgcttg	agcagacatc	2160
accatgcctg	aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	gttcgacagc	2220
gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	cttcgatgta	2280
ggagggcgtg	gatatgtcct	gcgggtaaat	agctgcgccg	atggtttcta	caaagatcgt	2340
tatgtttatc	ggcactttgc	ateggeegeg	ctcccgattc	cggaagtgct	tgacattggg	2400
gaattcagcg	agagcctgac	ctattgcatc	tcccgccgtg	cacagggtgt	cacgttgcaa	2460
gacctgcctg	aaaccgaact	gcccgctgtt	ctgcagccgg	tcgcggaggc	catggatgcg	2520
ategetgegg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaatc	2580
ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	tgtgtatcac	2640
tggcaaactg	tgatggacga	caccgtcagt	gegteegteg	cgcaggctct	: cgatgagctg	2700
atgctttggg	ccgaggactg	cecegaagte	: cggcacctcg	tgcacgcgga	tttcggctcc	2760

WO 2004/063359 PCT/EP2004/000099 231/357

aacaatgtcc	tgacggacaa	tggccgcata	acagcggtca	ttgactggag	cgaggcgatg	2820
ttcggggatt	cccaatacga	ggtcgccaac	atcttcttct	ggaggccgtg	gttggcttgt	2880
atggagcagc	agacgcgcta	cttcgagcgg	aggcatccgg	agcttgcagg	ategeegegg	2940
ctccgggcgt	atatgctccg	cattggtctt	gaccaactct	atcagagctt	ggttgacggc	3000
aatttcgatg	atgcagcttg	ggcgcagggt	cgatgcgacg	caatcgtccg	atccggagcc	3060
gggactgtcg	ggcgtacaca	aatcgcccgc	agaagcgcgg	ccgtctggac	cgatggctgt	3120
gtagaagtac	tcgccgatag	tggaaaccga	cgccccagca	ctcgtccgag	ggcaaaggaa	3180
tagagtagat	gccgaccgcg	ggatcgatcc	acttaacgtt	actgaaatca	tcaaacagct	3240
tgacgaatct	ggatataaga	tcgttggtgt	cgatgtcagc	tccggagttg	agacaaatgg	3300
tgttcaggat	ctcgataaga	tacgttcatt	tgtccaagca	gcaaagagtg	ccttctagtg	3360
atttaatagc	tccatgtcaa	caagaataaa	acgcgttttc	gggtttacct	cttccagata	3420
cageteatet	gcaatgcatt	aatgcattga	ctgcaaccta	gtaacgcctt	ncaggeteeg	3480
gcgaagagaa	gaatagctta	gcagagctat	tttcattttc	gggagacgag	atcaagcaga	3540
tcaacggtcg	tcaagagacc	tacgagactg	aggaatccgc	tettggetee	acgcgactat	3600
atatttgtct	ctaattgtac	tttgacatgc	tectettett	tactctgata	gcttgactat	3660
gaaaattccg	tcaccagcnc	ctgggttcgc	aaagataatt	gcatgtttct	tccttgaact	3720
ctcaagccta	caggacacac	attcatcgta	ggtataaacc	tcgaaatcan	ttcctactaa	3780
gatggtatac	aatagtaacc	atgcatggtt	gcctagtgaa	tgctccgtaa	cacccaatac	3840
gccggccgaa	actttttac	aactctccta	tgagtcgttt	acccagaatg	cacaggtaca	3900
cttgtttaga	ggtaatcctt	ctttctagct	agaagtcctc	gtgtactgtg	taagcgccca	3960
ctccacatct	: ccactcgacc	tgcaggcatg	caagcttgaa	ttcgagctcg	gtacccgggg	4020

WO 2004/063359 PCT/EP2004/000099 232/357

atctttcgac	actgaaatac	gtcgagcctg	ctccgcttgg	aagcggcgag	gagcctcgtc	4080
ctgtcacaac	taccaacatg	gagtacgata	agggccagtt	ccgccagete	attaagagcc	4140
agttcatggg	cgttggcatg	atggccgtca	tgcatctgta	cttcaagtac	accaacgctc	4200
ttctgatcca	gtcgatcatc	cgctgaaggc	gctttcgaat	ctggttaaga	tccacgtctt	4260
cgggaagcca	gcgactggtg	acctccageg	tccctttaag	gctgccaaca	gctttctcag	4320
ccagggccag	cccaagaccg	acaaggcctc	cctccagaac	gccgagaaga	actggagggg	4380
tggtgtcaag	gaggagtaag	ctccttattg	aagtcggagg	acggagcggt	gtcaagagga	4440
tattcttcga	ctctgtatta	tagataagat	gatgaggaat	tggaggtagc	atagcttcat	4500
ttggatttgc	tttccaggct	gagactctag	cttggagcat	agagggtcct	ttggctttca	4560
atattctcaa	gtatctcgag	tttgaactta	ttccctgtga	accttttatt	caccaatgag	4620
cattggaatg	aacatgaatc	tgaggactgc	aatcgccatg	aggttttcga	aatacatccg	4680
gatgtcgaag	gcttggggca	cctgcgttgg	ttgaatttag	aacgtggcac	tattgatcat	4740
ccgatagctc	tgcaaagggc	gttgcacaat	gcaagtcaaa	cgttgctagc	agttccaggt	4800
ggaatgttat	gatgagcatt	gtattaaatc	aggagatata	gcatgatctc	tagttagctc	4860
accacaaaag	tcagacggcg	taaccaaaag	tcacacaaca	caagctgtaa	ggatttcggc	4920
acggctacgg	aagacggaga	agccaccttc	agtggactcg	agtaccattt	aattctattt	4980
gtgtttgatc	gagacctaat	acageceeta	caacgaccat	caaagtcgta	tagctaccag	5040
tgaggaagtg	gactcaaatc	gacttcagca	acatctcctg	gataaacttt	aagcctaaac	5100
tatacagaat	aagataggtg	gagagettat	accgagetec	caaatctgtc	cagatcatgg	5160
ttgaccggtg	cctggatctt	cctatagaat	catccttatt	cgttgaccta	gctgattctg	5220
gagtgaccca	gagggtcatg	acttgagcct	aaaatccgcc	gcctccacca	tttgtagaaa	5280

WO 2004/063359 PCT/EP2004/000099 233/357

aatgtgacga actcgtgagc tctgtacag	t gaccggtgac	tctttctggc	atgcggagag	5340
acggacggac gcagagagaa gggctgagt	a ataagccact	ggccagacag	ctctggcggc	5400
tctgaggtgc agtggatgat tattaatcc	g ggaccggccg	ccctccgcc	ccgaagtgga	5460
aaggctggtg tgcccctcgt tgaccaaga	a tctattgcat	catcggagaa	tatggagctt	5520
catcgaatca ccggcagtaa gcgaaggag	a atgtgaagcc	aggggtgtat	agccgtcggc	5580
gaaatagcat gccattaacc taggtacag	a agtccaattg	cttccgatct	ggtaaaagat	5640
tcacgagata gtaccttctc cgaagtagg	t agagcgagta	cccggcgcgt	aagctcccta	5700
attggcccat ccggcatctg tagggcgtc	c aaatatcgtg	cctctcctgc	tttgcccggt	5760
gtatgaaacc ggaaaggccg ctcaggagc	t ggccagcggc	gcagaccggg	aacacaagct	5820
ggcagtcgac ccatccggtg ctctgcact	c gacctgctga	ggtccctcag	tccctggtag	5880
gcagetttge ceegtetgte egeceggtg	t gtcggcgggg	ttgacaaggt	cgttgcgtca	5940
gtccaacatt tgttgccata ttttcctgc	t ctccccacca	gctgctcttt	tettttetet	6000
ttetttteec atetteagta tatteatet	t cccatccaag	aacctttatt	tcccctaagt	6060
aagtactttg ctacatccat actccatcc	t teccatecet	tattcctttg	aacctttcag	6120
ttcgagcttt cccacttcat cgcagcttc	ga ctaacagcta	ccccgcttga	gcagacatca	6180
ccatgtcaat actcacttat ctggaattt	c atctctacta	ı tacactacct	gtccttgcgg	6240
cattgtgttg gctgctaaag ccgtttcac	ct cacagcaaga	caatctcaag	tataaatttt	6300
taatgttgat ggccgcctct accgcatcg	ga tttgggacaa	ı ttatatcgtt	tatcatcgcg	6360
cttggtggta ctgtcctact tgtgttgtg	gg ctgtcattgg	g ctatgtacct	ctagaagaat	6420
acatgttctt tatcatcatg actttaatg	ga ctgtcgcgtt	: ctcaaacttt	gttatgcgtt	6480
ggcacttgca tactttcttt attagacco	ca acacttcttç	g gaagcaaaca	ctattagtac	6540

WO 2004/063359 PCT/EP2004/000099 234/357

gccttgtgcc tgtttcagct ttattggcaa tcacttatca tgcttggcac ttgacactgc	6600
caaataaacc ttcattttat ggttcatgca tcctttggta tgcttgtcct gtgttggcta	6660
ttctttggct gggtgctggc gaatatatct tgcgtcgacc tgtggctgtc cttttgtcta	6720
ttgttatccc tagtgtatac ctatgttggg ctgatatcgt cgctattagt gctggcacat	6780
ggcatatttc tcttagaaca agcactggca aaatggtagt acccgattta cctgtagaag	6840
aatgcctgtt ttttactttg atcaacacag tcttggtttt tgctacctgt gctatagacc	6900
gcgctcaggc catcctccat gtgagcgcgc gtaatacgac tcactatagg gcgaattgga	6960
gctccaccgc ggtggcggcc gctctagaac tagtggatcc cccgggctgc aggaattcgg	7020
cacgagetae attteacaag eeegtgageg gtgeaagege tetgeeceae ateggeecae	7080
ctcctcatct ccatcggtca tttgctgcta ccacgatgct gtcgaagctg cagtcaatca	7140
gcgtcaaggc ccgccgcgtt gaactagccc gcgacatcac gcggcccaaa gtctgcctgc	7200
atgctcagcg gtgctcgtta gttcggctgc gagtggcagc accacagaca gaggaggcgc	7260
tgggaaccgt gcaggctgcc ggcgcgggcg atgagcacag cgccgatgta gcactccagc	7320
agettgaceg ggetategea gagegtegtg ceeggegeaa aegggageag etgteatace	7380
aggetgeege cattgeagea teaattggeg tgteaggeat tgeeatette geeacetace	7440
tgagatttgc catgcacatg accgtgggcg gcgcagtgcc atggggtgaa gtggctggca	7500
ctctcctctt ggtggttggt ggcgcgctcg gcatggagat gtatgcccgc tatgcacaca	7560
aagccatctg gcatgagtcg cetetggget ggetgetgea caagageeae cacacacete	7620
gcactggacc ctttgaagcc aacgacttgt ttgcaatcat caatggactg cccgccatgc	7680
tectgtgtae etttggette tggetgeeca aegteetggg ggeggeetge tttggagegg	7740
ggctgggcat cacgctatac ggcatggcat atatgtttgt acacgatggc ctggtgcaca	7800

WO 2004/063359 PCT/EP2004/000099 235/357

ggcgctttcc	caccgggccc	atcgctggcc	tgccctacat	gaagcgcctg	acagtggccc	7860
accagctaca	ccacagcggc	aagtacggtg	gegegeeetg	gggtatgttc	ttgggtccac	7920
aggagctgca	gcacattcca	ggtgcggcgg	aggaggtgga	gcgactggtc	ctggaactgg	798 0 .
actggtccaa	gcgggctcag	gccatcctcc	atctgtacaa	atcatctgtt	caaaatcaaa	8040
accctaaaca	agccatttcc	cttttccagc	atgtcaaaga	gctagcatgg	gccttctgtc	8100
ttcctgacca	aatgctcaac	aatgaattgt	ttgatgatct	tactatcagc	tgggatattt	8160
tacgtaaagc	ctcaaagtca	ttctatactg	catctgccgt	ttttccaagt	tatgtacgtc	8220
aagacttggg	tgttctctat	gctttctgca	gagctaccga	tgacctgtgc	gatgatgaat	8280
ccaaatctgt	tcaagaaaga	agagaccaat	tagatcttac	tcgacaattt	gttcgtgatc	8340
tctttagcca	aaagaccagt	gcgcctattg	tgattgattg	ggaattgtat	caaaaccaac	8400
ttcctgcttc	ttgtatatca	gcctttagag	cctttactcg	ccttcgccat	gtccttgaag	8460
tagaccctgt	agaagaacta	ttagatggtt	acaaatggga	tcttgagcgt	cgtcctatcc	8520
ttgatgaaca	agacttggag	gcatactctg	cttgtgtggc	cagtagtgtg	ggtgaaatgt	8580
gcacacgtgt	gattcttgct	caagaccaaa	aggaaaatga	tgcttggata	attgaccgtg	8640
cacgtgagat	ggggctggtg	ctacaatacg	ttaacattgc	tegagacatt	gtgactgata	8700
gcgagactct	gggtcgatgt	tatctgcctc	aacaatggct	tagaaaagaa	gaaacagaac	8760
aaatacagca	aggcaacgcc	cgtagcctag	gtgatcaaag	actgttgggc	ttgtctctga	8820
agcttgtagg	aaaggcagac	gctatcatgg	tgagagctaa	gaagggcatt	gacaagttgc	8880
cggcaaactg	tcaaggcggt	gtacgagctg	cttgccaagt	atatgctgca	attggatctg	8940
tactcaagca	gcagaagaca	acatatccta	caagagctca	tctaaaagga	agcgaacgtg	9000
ccaagattgc	tctgttgagt	gtatacaacc	tctatcaatc	tgaagacaag	cctgtggctc	9060

WO 2004/063359 PCT/EP2004/000099 236/357

tccgtcaagc	tagaaagatt	aagagttttt	ttgttgatta	gtgaatttt	gttttattta	9120
tgtctgatag	ttcaataaag	agacaacaca	tacaatataa	aatcattgtc	tttaaatgtt	9180
aatttagtag	agtgtaaagc	ctgcattttt	tttgtacgca	taaacaatga	gttcaccccg	9240
cttctggttt	ttaaataatt	atgtcaaact	agggaaaatt	ctttttttc	tcttcgttct	9300
ttttttggct	tgttgtggag	tcacaggett	gtcttcagat	tgatagaggt	tgtatacact	9360
caacagagca	atcttggcac	gttcgcttcc	ttttagatga	gctcttgtag	gatatgttgt	9420
cttctgctgc	ttgagtacag	atccaattgc	agcatatact	tggcaagcag	ctcgtacacc	9480
gccttgacag	tttgccggca	acttgtcaat	gcccttctta	gctctcacca	tgatagcgtc	9540
tgcctttcct	acaagcttgg	cgtaatcatg	gtcatagctg	tttcctgtgt	gaaattgtta	9600
teegeteaca	attccacaca	acatacgagc	cggaagcata	aagtgtaaag	cctggggtgc	9660
ctaatgagtg	agctaactca	cattaattgc	gttgcgctca	ctgcccgctt	tccagtcggg	9720
aaacctgtcg	tgccagctgc	attaatgaat	cggccaacgc	gcggggagag	gcggtttgcg	9780
tattgggcca	aagacaaaag	ggcgacattc	aaccgattga	gggagggaag	gtaaatattg	9840
acggaaatta	ttcattaaag	gtgaattatc	accgtcaccg	acttgagcca	tttgggaatt	9900
agagccagca	aaatcaccag	tagcaccatt	accattagca	aggccggaaa	cgtcaccaat	9960
gaaaccatcg	atagcagcac	cgtaatcagt	agcgacagaa	tcaagtttgc	ctttagcgtc	10020
agactgtagc	gcgttttcat	cggcattttc	ggtcatagcc	cccttattag	cgtttgccat	10080
cttttcataa	tcaaaatcac	: cggaaccaga	gccaccaccg	gaaccgcctc	cctcagagcc	10140
gccaccctca	ı gaaccgccac	: cctcagagec	accaccctca	gageegeeac	: cagaaccacc	10200
accagagecg	g ccgccagcat	: tgacaggagg	g cccgatctag	taacatagat	gacaccgcgc	10260
gcgataattt	atcctagttt	gegegetata	ttttgttttc	: tatcgcgtat	: taaatgtata	10320

WO 2004/063359 PCT/EP2004/000099 237/357

attgcgggac tctaatcata aaaacccatc tcataaataa cgtcatgcat tacatgttaa 10380 ttattacatg cttaacgtaa ttcaacagaa attatatgat aatcatcgca agaccggcaa 10440 caggattcaa tottaagaaa otttattgoo aaatgtttga acgatcgggg atcatccggg 10500 tetgtggegg gaactecaeg aaaatateeg aacgeageaa gatategegg tgeatetegg 10560 tcttgcctgg gcagtcgccg ccgacgccgt tgatgtggac gccgggcccg atcatattgt 10620 cgctcaggat cgtggcgttg tgcttgtcgg ccgttgctgt cgtaatgata tcggcacctt 10680 cgaccgcctg ttccgcagag atcccgtggg cgaagaactc cagcatgaga tccccgcgct 10740 ggaggatcat ccagccggcg tcccggaaaa cgattccgaa gcccaacctt tcatagaagg 10800 cggcggtgga atcgaaatct cgtgatggca ggttgggcgt cgcttggtcg gtcatttcga 10860 accccagagt cccgctcaga agaactcgtc aagaaggcga tagaaggcga tgcgctgcga 10920 atcgggagcg gcgataccgt aaagcacgag gaagcggtca gcccattcgc cgccaagctc 10980 ttcagcaata tcacgggtag ccaacgctat gtcctgatag cggtccgcca cacccagccg 11040 gccacagtcg atgaatccag aaaagcggcc attttccacc atgatattcg gcaagcaggc 11100 atcgccatgg gtcacgacga gatcatcgcc gtcgggcatg cgcgccttga gcctggcgaa 11160 cagttegget ggegegagee cetgatgete ttegtecaga teateetgat egacaagace 11220 ggcttccatc cgagtacgtg ctcgctcgat gcgatgtttc gcttggtggt cgaatgggca 11280 ggtageegga teaagegtat geageegeeg cattgeatea geeatgatgg ataetttete 11340 ggcaggagca aggtgagatg acaggagatc ctgccccggc acttcgccca atagcagcca 11400 gtcccttccc gcttcagtga caacgtcgag cacagctgcg caaggaacgc ccgtcgtggc 11460 cagccacgat agccgcgctg cetegteetg cagtteatte agggcaccgg acaggteggt 11520 cttgacaaaa agaaccgggc gcccctgcgc tgacagccgg aacacggcgg catcagagca 11580

WO 2004/063359 PCT/EP2004/000099 238/357

gccgattgtc	tgttgtgccc	agtcatagcc	gaatagcctc	tccacccaag	cggccggaga	11640
acctgcgtgc	aatccatctt	gttcaatcat	gcgaaacgat	ccagatccgg	tgcagattat	11700
ttggattgag	agtgaatatg	agactctaat	tggataccga	ggggaattta	tggaacgtca	11760
gtggagcatt	tttgacaaga	aatatttgct	agctgatagt	gaccttaggc	gacttttgaa	11820
cgcgcaataa	tggtttctga	cgtatgtgct	tagctcatta	aactccagaa	acccgcggct	11880
gagtggctcc	ttcaacgttg	eggttetgte	agttccaaac	gtaaaacggc	ttgtcccgcg	11940
tcatcggcgg	gggtcataac	gtgactccct	taattctccg	ctcatgatca	gattgtcgtt	12000
tecegeette	agtttaaact	atcagtgttt	gacaggatat	attggcgggt	aaacctaaga	12060
gaaaagagcg	tttattagaa	taatcggata	tttaaaaggg	cgtgaaaagg	tttatccgtt	12120
cgtccatttg	tatgtgcatg	ccaaccacag	ggttccccag	atctggcgcc	ggccagcgag	12180
acgagcaaga	ttggccgccg	cccgaaacga	teegacageg	cgcccagcac	aggtgcgcag	12240
gcaaattgca	ccaacgcata	cagcgccagc	agaatgccat	agtgggcggt	gacgtcgttc	12300
gagtgaacca	gategegeag	gaggcccggc	agcaccggca	taatcaggcc	gatgccgaca	12360
gcgtcgagcg	cgacagtgct	cagaattacg	atcaggggta	tgttgggttt	cacgtctggc	12420
ctccggacca	gcctccgctg	gtccgattga	acgcgcggat	tctttatcac	tgataagttg	12480
gtggacatat	tatgtttatc	agtgataaag	tgtcaagcat	gacaaagttg	cagccgaata	12540
cagtgatccg	tgccgccctg	gacctgttga	acgaggtcgg	cgtagacggt	ctgacgacac	12600
gcaaactggc	ggaacggttg	ggggttcagc	agccggcgct	ttactggcac	ttcaggaaca	12660
agcgggcgct	gctcgacgca	ctggccgaag	ccatgctggc	ggagaatcat	acgcattcgg	12720
tgccgagagc	cgacgacgac	tggcgctcat	ttctgatcgg	gaatgecege	agcttcaggc	12780
aggegetget	cgcctaccgc	gatggcgcgc	gcatccatgo	cggcacgcga	ccgggcgcac	12840

WO 2004/063359 PCT/EP2004/000099 239/357

cgcagatgga	aacggccgac	gcgcagcttc	gcttcctctg	cgaggcgggt	ttttcggccg	12900
gggacgccgt	caatgcgctg	atgacaatca	gctacttcac	tgttggggcc	gtgcttgagg	12960
agcaggccgg	cgacagcgat	gccggcgagc	gcggcggcac	cgttgaacag	gctccgctct	13020
cgccgctgtt	gcgggccgcg	atagacgcct	tcgacgaagc	cggtccggac	gcagcgttcg	13080
agcagggact	cgcggtgatt	gtcgatggat	tggcgaaaag	gaggctcgtt	gtcaggaacg	13140
ttgaaggacc	gagaaagggt	gacgattgat	caggaccgct	gccggagcgc	aacccactca	13200
ctacagcaga	gccatgtaga	caacatcccc	tcccctttc	caccgcgtca	gacgcccgta	13260
gcagcccgct	acgggctttt	tcatgccctg	ccctagcgtc	caagcctcac	ggccgcgctc	13320
ggcctctctg	gcggccttct	ggcgctcttc	cgcttcctcg	ctcactgact	cgctgcgctc	13380
ggtcgttcgg	ctgcggcgag	cggtatcagc	tcactcaaag	gcggtaatac	ggttatccac	13440
agaatcaggg	gataacgcag	gaaagaacat	gtgagcaaaa	ggccagcaaa	aggccaggaa	13500
ccgtaaaaag	gccgcgttgc	tggcgttttt	ccataggctc	egececectg	acgagcatca	13560
caaaaatcga	cgctcaagtc	agaggtggcg	aaacccgaca	ggactataaa	gataccaggc	13620
gtttccccct	ggaagctccc	tegtgegete	tectgtteeg	accctgccgc	ttaccggata	13680
cctgtccgcc	tttctccctt	cgggaagcgt	ggcgcttttc	cgctgcataa	ccctgcttcg	13740
gggtcattat	agcgatttt	teggtatate	catccttttt	cgcacgatat	acaggatttt	13800
gccaaagggt	tcgtgtagac	tttccttggt	gtatccaacg	gcgtcagccg	ggcaggatag	13860
gtgaagtagg	cccacccgcg	agcgggtgtt	ccttcttcac	tgtcccttat	tegeacetgg	13920
cggtgctcaa	cgggaatcct	gctctgcgag	getggeegge	taccgccggc	gtaacagatg	13980
agggcaagcg	gatggctgat	gaaaccaagc	caaccaggaa	gggcagccca	cctatcaagg	14040
tgtactgcct	tccagacgaa	cgaagagcga	ttgaggaaaa	ggcggcggcg	gccggcatga	14100

WO 2004/063359 PCT/EP2004/000099 240/357

geetgtegge	ctacctgctg	geegteggee	agggctacaa	aatcacgggc	gtcgtggact	14160
atgagcacgt	ccgcgagctg	gcccgcatca	atggcgacct	gggccgcctg	ggcggcctgc	14220
tgaaactctg	gctcaccgac	gacccgcgca	cggcgcggtt	cggtgatgcc	acgatcctcg	14280
ccctgctggc	gaagatcgaa	gagaagcagg	acgagcttgg	caaggtcatg	atgggcgtgg	14340
tccgcccgag	ggcagagcca	tgacttttt	agccgctaaa	acggccgggg	ggtgcgcgtg	14400
attgccaagc	acgtccccat	gcgctccatc	aagaagagcg	acttcgcgga	gctggtgaag	14460
tacatcaccg	acgagcaagg	caagaccgag	cgcctttgcg	acgctcaccg	ggctggttgc	14520
cctcgccgct	gggctggcgg	ccgtctatgg	ccctgcaaac	gcgccagaaa	cgccgtcgaa	14580
gccgtgtgcg	agacaccgcg	gccgccggcg	ttgtggatac	ctcgcggaaa	acttggccct	14640
cactgacaga	tgaggggcgg	acgttgacac	ttgaggggcc	gactcacccg	gegeggegtt	14700
gacagatgag	gggcaggctc	gatttcggcc	ggcgacgtgg	agctggccag	cctcgcaaat	14760
cggcgaaaac	gcctgatttt	acgcgagttt	cccacagatg	atgtggacaa	gcctggggat	14820
aagtgccctg	cggtattgac	acttgagggg	egegaetaet	gacagatgag	gggcgcgatc	14880
cttgacactt	gaggggcaga	gtgctgacag	atgaggggcg	cacctattga	catttgaggg	14940
gctgtccaca	ggcagaaaat	ccagcatttg	caagggtttc	cgcccgtttt	teggecaceg	15000
ctaacctgtc	ttttaacctg	cttttaaacc	aatatttata	aaccttgttt	ttaaccaggg	15060
ctgcgccctg	tgcgcgtgac	cgcgcacgcc	gaaggggggt	gcccccctt	ctcgaaccct	15120
cccggcccgc	taacgcgggc	ctcccatccc	cccaggggct	gegeeeeteg	gccgcgaacg	15180
gcctcacccc	aaaaatggca	gegetggeag	tccttgccat	tgccgggatc	ggggcagtaa	15240
cgggatgggc	gatcagcccg	agegegaege	ccggaagcat	tgacgtgccg	caggtgctgg	15300
catcgacatt	cagcgaccag	gtgccgggca	gtgagggcgg	cggcctgggt	ggcggcctgc	15360

WO 2004/063359 PCT/EP2004/000099 241/357

ccttcacttc	ggccgtcggg	gcattcacgg	acttcatggc	ggggccggca	atttttacct	15420
tgggcattct	tggcatagtg	gtcgcgggtg	ccgtgctcgt	gttcgggggt	gcgataaacc	15480
cagcgaacca	tttgaggtga	taggtaagat	tataccgagg	tatgaaaacg	agaattggac	15540
ctttacagaa	ttactctatg	aagcgccata	tttaaaaagc	tacçaagacg	aagaggatga	15600
agaggatgag	gaggcagatt	gccttgaata	tattgacaat	actgataaga	taatatatct	15660
tttatataga	agatatcgcc	gtatgtaagg	atttcagggg	gcaaggcata	ggcagcgcgc	15720
ttatcaatat	atctatagaa	tgggcaaagc	ataaaaactt	gcatggacta	atgcttgaaa	15780
cccaggacaa	taaccttata	gcttgtaaat	tctatcataa	ttgggtaatg	actccaactt	15840
attgatagtg	tttatgttc	agataatgcc	cgatgacttt	gtcatgcagc	tccaccgatt	15900
ttgagaacga	cagcgacttc	cgtcccagcc	gtgccaggtg	ctgcctcaga	ttcaggttat	15960
gccgctcaat	tcgctgcgta	tatcgcttgc	tgattacgtg	cagetttece	ttcaggcggg	16020
attcatacag	cggccagcca	tccgtcatcc	atatcaccac	gtcaaagggt	gacagcaggc	16080
tcataagacg	ccccagcgtc	gccatagtgc	gttcaccgaa	tacgtgcgca	acaaccgtct	16140
tccggagact	gtcatacgcg	taaaacagcc	agcgctggcg	cgatttagcc	ccgacatagc	16200
cccactgttc	gtccatttcc	gcgcagacga	tgacgtcact	gcccggctgt	atgcgcgagg	16260
ttaccgactg	cggcctgagt	tttttaagtg	acgtaaaatc	gtgttgaggc	caacgcccat	16320
aatgcgggct	gttgcccggc	atccaacgcc	attcatggcc	atatcaatga	ttttctggtg	16380
cgtaccgggt	tgagaagcgg	tgtaagtgaa	ctgcagttgc	catgttttac	ggcagtgaga	16440
gcagagatag	cgctgatgtc	cggcggtgct	tttgccgtta	cgcaccaccc	cgtcagtagc	16500
tgaacaggag	ggacagctga	tagacacaga	agccactgga	gcacctcaaa	aacaccatca	16560
tacactaaat	cagtaagttg	gcagcatcac	ccataattgt	ggtttcaaaa	teggeteegt	16620

WO 2004/063359 PCT/EP2004/000099 242/357

cgatactatg	ttatacgcca	actttgaaaa	caactttgaa	aaagctgttt	tctggtattt	16680
aaggttttag	aatgcaagga	acagtgaatt	ggagttcgtc	ttgttataat	tagettettg	16740
gggtatcttt	aaatactgta	gaaaagagga	aggaaataat	aaatggctaa	aatgagaata	16800
tcaccggaat	tgaaaaaact	gatcgaaaaa	taccgctgcg	taaaagatac	ggaaggaatg	16860
tctcctgcta	aggtatataa	gctggtggga	gaaaatgaaa	acctatattt	aaaaatgacg	16920
gacagccggt	ataaagggac	cacctatgat	gtggaacggg	aaaaggacat	gatgctatgg	16980
ctggaaggaa	agctgcctgt	tccaaaggtc	ctgcactttg	aacggcatga	tggctggagc	17040
aatctgctca	tgagtgaggc	cgatggcgtc	ctttgctcgg	aagagtatga	agatgaacaa	17100
agccctgaaa	agattatcga	gctgtatgcg	gagtgcatca	ggctctttca	ctccatcgac	17160
atatcggatt	gtccctatac	gaatagctta	gacagccgct	tagccgaatt	ggattactta	17220
ctgaataacg	atctggccga	tgtggattgc	gaaaactggg	aagaagacac	tccatttaaa	17280
gatccgcgcg	agctgtatga	ttttttaaag	acggaaaagc	ccgaagagga	acttgtcttt	17340
tcccacggcg	acctgggaga	cagcaacatc	tttgtgaaag	atggcaaagt	aagtggcttt	17400
attgatcttg	ggagaagcgg	cagggcggac	aagtggtatg	acattgcctt	ctgcgtccgg	17460
tcgatcaggg	aggatatcgg	ggaagaacag	tatgtcgagc	tattttttga	cttactgggg	17520
atcaagcctg	attgggagaa	aataaaatat	tatattttac	tggatgaatt	gttttagtac	17580
ctagatgtgg	cgcaacgatg	ccggcgacae	gcaggagcgc	accgacttct	tccgcatcaa	17640
gtgttttggc	tctcaggccg	aggcccacgg	caagtatttg	ggcaaggggt	cgctggtatt	17700
cgtgcagggc	aagattcgga	ataccaagta	cgagaaggac	ggccagacgg	tctacgggac	17760
cgacttcatt	gccgataagg	tggattatct	ggacaccaag	gcaccaggcg	ggtcaaatca	17820
ggaataaggg	cacattgccc	cggcgtgagt	cggggcaatc	ccgcaaggag	ggtgaatgaa	17880

WO 2004/063359 PCT/EP2004/000099 243/357

tcggacgttt	gaccggaagg	catacaggca	agaactgatc	gacgcggggt	tttccgccga	17940
ggatgccgaa	accatcgcaa	gccgcaccgt	catgcgtgcg	ccccgcgaaa	ccttccagtc	18000
cgtcggctcg	atggtccagc	aagctacggc	caagatcgag	cgcgacagcg	tgcaactggc	18060
teceetgee	ctgcccgcgc	categgeege	cgtggagcgt	tegegtegte	tcgaacagga	18120
ggcggcaggt	ttggcgaagt	cgatgaccat	cgacacgcga	ggaactatga	cgaccaagaa	18180
gcgaaaaacc	gccggcgagg	acctggcaaa	acaggtcagc	gaggccaagc	aggccgcgtt	18240
gctgaaacac	acgaagcagc	agatcaagga	aatgcagctt	tecttgtteg	atattgcgcc	18300
gtggccggac	acgatgcgag	c <u>g</u> atgccaaa	cgacacggcc	cgctctgccc	tgttcaccac	18360
gcgcaacaag	aaaatcccgc	gcgaggcgct	gcaaaacaag	gtcattttcc	acgtcaacaa	18420
ggacgtgaag	atcacctaca	ccggcgtcga	gctgcgggcc	gacgatgacg	aactggtgtg	18480
gcagcaggtg	ttggagtacg	cgaagcgcac	ccctatcggc	gagccgatca	ccttcacgtt	18540
ctacgagctt	tgccaggacc	tgggctggtc	gatcaatggc	cggtattaca	cgaaggccga	18600
ggaatgcctg	tcgcgcctac	aggcgacggc	gatgggcttc	acgtccgacc	gcgttgggca	18660
cctggaatcg	gtgtcgctgc	tgcaccgctt	cegegteetg	gaccgtggca	agaaaacgtc	18720
ccgttgccag	gtcctgatcg	acgaggaaat	cgtcgtgctg	tttgctggcg	accactacac	18780
gaaattcata	tgggagaagt	accgcaagct	gtcgccgacg	gcccgacgga	tgttcgacta	18840
tttcagctcg	caccgggagc	cgtacccgct	caagctggaa	accttccgcc	tcatgtgcgg	18900
atcggattcc	accegegtga	agaagtggcg	cgagcaggtc	ggcgaagcct	gcgaagagtt	18960
gcgaggcagc	ggcctggtgg	aacacgcctg	ggtcaatgat	gacctggtgc	attgcaaacg	19020
ctagggcctt	gtggggtcag	ttccggctgg	gggttcagca	gccagcgctt	tactggcatt	19080
tcaggaacaa	gcgggcactg	ctcgacgcac	ttgcttcgct	cagtatcgct	cgggacgcac	19140

WO 2004/063359 PCT/EP2004/000099 244/357

ggcgcgctct acgaactgcc	gataaacaga	ggattaaaat	tgacaattgt	gattaaggct	19200
cagattcgac ggcttggagc	ggccgacgtg	caggatttcc	gcgagatccg	attgtcggcc	19260
ctgaagaaag ctccagagat	gttcgggtcc	gtttacgagc	acgaggagaa	aaagcccatg	19320
gaggegtteg etgaaeggtt	gcgagatgcc	gtggcattcg	gegeetacat	cgacggcgag	19380
atcattgggc tgtcggtctt	caaacaggag	gacggcccca	aggacgctca	caaggcgcat	19440
ctgtccggcg ttttcgtgga	gcccgaacag	cgaggccgag	gggtcgccgg	tatgctgctg	19500
cgggcgttgc cggcgggttt	attgctcgtg	atgatcgtcc	gacagattcc	aacgggaatc	19560
tggtggatgc gcatcttcat	cctcggcgca	cttaatattt	cgctattctg	gagcttgttg	19620
tttatttcgg tctaccgcct	gccgggcggg	gtcgcggcga		tgtgcagccg	19680
ctgatggtcg tgttcatctc	tgccgctctg	ctaggtagec	cgatacgatt	gatggcggtc	19740
ctgggggcta tttgcggaac	tgegggegtg	gegetgttgg	tgttgacacc	aaacgcagcg	19800
ctagatectg teggegtege	agegggeetg	gcgggggcgg	tttccatggc	gttcggaacc	19860
gtgctgaccc gcaagtggca	acctcccgtg	cctctgctca	cctttaccgc	ctggcaactg	19920
gcggccggag gacttctgct	cgttccagta	gctttagtgt	ttgatccgcc	aatcccgatg	19980
cctacaggaa ccaatgttct	cggcctggcg	tggctcggcc	tgatcggagc	gggtttaacc	20040
tacttccttt ggttccgggg	gatctcgcga	ctcgaaccta	cagttgtttc	cttactgggc	20100
tttctcagcc ccagatctgg	ggtcgatcag	ccggggatgc	atcaggccga	cagtcggaac	20160
ttegggteee egaeetgtad	: cattcggtga	gcaatggata	ggggagttga	tatcgtcaac	20220
gttcacttct aaagaaatag	g cgccactcag	cttcctcago	ggctttatco	agcgatttcc	20280
tattatgtcg gcatagttct	: caagatcgac	agcctgtcac	ggttaagcga	gaaatgaata	20340
agaaggctga taattcgga	ctctgcgagg	gagatgatat	ttgatcacag	gcagcaacgc	20400

tctgtcatcg	ttacaatcaa	catgctaccc	tccgcgagat	catccgtgtt	tcaaacccgg	20460
cagcttagtt	gccgttcttc	cgaatagcat	cggtaacatg	agcaaagtct	geegeettae	20520
aacggctctc	ccgctgacgc	cgtcccggac	tgatgggctg	cctgtatcga	gtggtgattt	20580
tgtgccgagc	tgccggtcgg	ggagctgttg	gctggctggt	ggcaggatat	attgtggtgt	20640
aaacaaattg	acgcttagac	aacttaataa	cacattgcgg	acgtttttaa	tgtactgggg	20700
tggtttttct	tttcaccagt	gagacgggca	acagctgatt	gcccttcacc	gcctggccct	20760
gagagagttg	cagcaagcgg	tccacgctgg	tttgccccag	caggcgaaaa ·	tcctgtttga	20820
tggtggttcc	gaaatcggca	aaatccctta	taaatcaaaa	gaatagcccg	agatagggtt	20880
gagtgttgtt	ccagtttgga	acaagagtcc	actattaaag	aacgtggact	ccaacgtcaa	20940
agggcgaaaa	accgtctatc	agggcgatgg	cccactacgt	gaaccatcac	ccaaatcaag	21000
ttttttgggg	tcgaggtgcc	gtaaagcact	aaatcggaac	cctaaaggga	gcccccgatt	21060
tagagcttga	cggggaaagc	cggcgaacgt	ggcgagaaag	gaagggaaga	aagcgaaagg	21120
agcgggcgcc	attcaggctg	cgcaactgtt	gggaagggcg	atcggtgcgg	gcctcttcgc	21180
tattacgcca	gctggcgaaa	gggggatgtg	ctgcaaggcg	attaagttgg	gtaacgccag	21240
ggttttccca	gtcacgacgt	tgtaaaacga	cggccagtga	attcgagctc	ggtacccggg	21300

<210> 47

<211> 17756

<212> DNA

<213> Artificial

<220>

<223> Plasmid

246/357 <221> misc_feature <222> (10264)..(10264) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10472)..(10472) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10563)..(10563) <223> n is a, c, g, or t <400> 47 ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60 120 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 180 ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240 cagcetegea aateggegaa aacgeetgat tttaegegag ttteecacag atgatgtgga 300 caagectggg gataagtgee etgeggtatt gacaettgag gggegegaet aetgacagat 360 420 gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 480 tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 540 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 600 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc cttctcgaac cctcccggec cgctaacgcg ggcctcccat cccccaggg gctgcgcccc 660 teggeegega acggeeteac eccaaaaatg geagegetgg eagteettge eattgeeggg 720 780 atcggggcag taacgggatg ggcgatcagc ccgagcgcga cgcccggaag cattgacgtg

ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg

840

WO 2004/063359 PCT/EP2004/000099 247/357

ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaattttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg	gacctttaca	gaattactct	atgaagcgcc	atatttaaaa	agctaccaag	1080
acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	aatactgata	1140
agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg	cgcttatcaa	tatatctata	gaatgggcaa	agcataaaaa	cttgcatgga	1260
ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgccgctc	aattcgctgc	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg	tcttccggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat	agccccactg	ttcgtccatt	teegegeaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg	aggttaccga	ctgcggcctg	agtttttaa	gtgacgtaaa	atcgtgttga	1800
ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg	agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100

WO 2004/063359 PCT/EP2004/000099 248/357

aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagc	tg 2160
ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgtt	at 2220
aattagcttc ttggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatg	gc 2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaa	ga 2340
tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaaccta	ta 2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaag	ga 2460
catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact ttgaacgg	ca 2520
tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct cggaagag	ta 2580
tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca tcaggctc	tt 2640
tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc gcttagcc	ga 2700
attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact gggaagaa	ıga 2760
cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa agcccgaa	ıga 2820
ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga aagatggc	aa 2880
agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggt atgacatt	.gc 2940
cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg agctattt	tt 3000
tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt tactggat	ga 3060
attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccga	act 3120
tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaa	agg 3180
ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggcca	aga 3240
cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcacc	cag 3300
gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca atcccgc	aag 3360

WO 2004/063359 PCT/EP2004/000099 249/357

gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
gugggeguue	9		-		55	
ggttttccgc	cgaggatgcc	gaaaccatcg	caagccgcac	cgtcatgcgt	gegeeeegeg	3480
aaaccttcca	gtccgtcggc	tegatggtee	agcaagctac	ggccaagatc	gagcgcgaca	3540
gcgtgcaact	ggctccccct	gccctgcccg	cgccatcggc	cgccgtggag	cgttcgcgtc	3600
gtctcgaaca	ggaggcggca	ggtttggcga	agtcgatgac	categacaeg	cgaggaacta	3660
tgacgaccaa	gaagcgaaaa	accgccggcg	aggacctggc	aaaacaggtc	agcgaggcca	3720
agcaggccgc	gttgctgaaa	cacacgaagc	agcagatcaa	ggaaatgcag	ctttccttgt	3780
tcgatattgc	gccgtggccg	gacacgatgc	gagcgatgcc	aaacgacacg	gecegetetg	3840
ccctgttcac	cacgcgcaac	aagaaaatcc	cgcgcgaggc	gctgcaaaac	aaggtcattt	3900
tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctatc	ggcgagccga	4020
tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggc	cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg	gcacctggaa	teggtgtege.	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200
gcaagaaaac	gtcccgttgc	caggtcctga	tcgacgagga	aatcgtcgtg	ctgtttgctg	4260
gcgaccacta	cacgaaattc	atatgggaga	agtaccgcaa	gctgtcgccg	acggcccgac	4320
ggatgttcga	ctatttcagc	tegcaceggg	agccgtaccc	gctcaagctg	gaaaccttcc	4380
gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
cctgcgaaga	gttgcgaggc	agcggcctgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca [.]	gcagccagcg	4560
ctttactggc	atttcaggaa	caagegggea	ctgctcgacg	cacttgcttc	gctcagtatc	4620

WO 2004/063359 PCT/EP2004/000099 250/357

	•		•			
gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
tgtgattaag	gctcagattc	gacggcttgg	ageggeegae	gtgcaggatt	tccgcgagat	4740
ccgattgtcg	gccctgaaga	aageteeaga	gatgttcggg	tccgtttacg	agcacgagga	4800
gaaaaagccc	atggaggcgt	tcgctgaacg	gttgcgagat	gccgtggcat	teggegeeta	4860
catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagcgaggcc	gaggggtcgc	4980
cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	tccgacagat	5040
tccaacggga	atctggtgga	tgcgcatctt	catcctcggc	gcacttaata	tttcgctatt	5100
ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160
cgctgtgcag	ccgctgatgg	tcgtgttcat	ctctgccgct	ctgctaggta	gcccgatacg .	5220
attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	tggtgttgac	5280
accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	cggtttccat	5340
ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	tcacctttac	5400
cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	tgtttgatcc	5460
gccaatcccg	atgcctacag	gaaccaatgt	tctcggcctg	gcgtggctcg	gcctgatcgg	5520
agcgggttta	acctacttcc	tttggttccg	ggggatctcg	cgactcgaac	ctacagttgt	5580
ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	tgcatcaggc	5640
cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	ataggggagt	5700
tgatatcgtc	aacgttcact	tctaaagaaa	tagegeeact	cagetteete	ageggettta	5760
tccagcgatt	tcctattatg	teggeatagt	tctcaagatc	gacagcctgt	cacggttaag	5820
cgagaaatga	ataagaaggo	: tgataattcg	gatctctgcg	agggagatga	tatttgatca	5880

WO 2004/063359 PCT/EP2004/000099 251/357

caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	gatcatccgt	5940
gtttcaaacc	cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	atgagcaaag	6000
tctgccgcct	tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	ctgcctgtat	6060
cgagtggtga	ttttgtgccg	agctgccggt	cggggagctg	ttggctggct	ggtggcagga	6120
tatattgtgg	tgtaaacaaa	ttgacgctta	gacaacttaa	taacacattg	cggacgtttt	6180
taatgtactg	gggtggtttt	tcttttcacc	agtgagacgg	gcaacagctg	attgcccttc	6240
accgcctggc	cctgagagag	ttgcagcaag	cggtccacgc	tggtttgccc	cagcaggcga	6300
aaatcctgtt	tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	6360
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480
cacccaaatc	aagttttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	6540
ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600
agaaagcgaa	aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt	cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720
tgggtaacgc	cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780
ctcggtaccc	ggggatcttt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840
cgaggagcct	cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900
gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac	gctcttctga	tccagtcgat	catccgctga	aggcgctttc	gaatctggtt	7020
aagatccacg	tcttcgggaa	gccagcgact	ggtgacetec	agcgtccctt	taaggctgcc	7080
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140

WO 2004/063359 PCT/EP2004/000099 252/357

aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag	aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320
tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
togaaataca	tccggatgtc	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560
tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaacatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ggtggagagc	ttataccgag	ctcccaaatc	7920
tgtccagatc	atggttgacc	ggtgcctgga	tcttcctata	gaatcatcct	tattcgttga	7980
cctagctgat	tctggagtga	cccagagggt	catgacttga	gcctaaaatc	cgccgcctcc	8040
accatttgta	. gaaaaatgtg	acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100
tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataago	: cactggccag	8160
acagetetgg	g cggctctgag	gtgcagtgga	tgattattaa	teegggaeeg	geegeeeete	8220
cgccccgaag	g tggaaaggct	ggtgtgcccc	: tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	a gcttcatcga	atcaccggca	ı gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatagccgt	cggcgaaata	gcatgccatt	: aacctaggta	ı cagaagtcca	attgcttccg	8400

WO 2004/063359 PCT/EP2004/000099 253/357

atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtacccggc	8460
gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc	8520
ctgctttgcc cggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac	8580
cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc	8640
tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca	8700
aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct	8760
cttttcttt ctctttcttt tcccatcttc agtatattca tcttcccatc caagaacctt	8820
tatttcccct aagtaagtac tttgctacat ccatactcca tccttcccat cccttattcc	8880
tttgaacctt tcagttcgag ctttcccact tcatcgcagc ttgactaaca gctacccgc	8940
ttgagcagac atcaccatgc ctgaactcac cgcgacgtct gtcgagaagt ttctgatcga	9000
aaagttcgac agcgtctccg acctgatgca gctctcggag ggcgaagaat ctcgtgcttt	9060
cagcttcgat gtaggagggc gtggatatgt cctgcgggta aatagctgcg ccgatggttt	9120
ctacaaagat cgttatgttt atcggcactt tgcatcggcc gcgctcccga ttccggaagt	9180
gcttgacatt ggggaattca gcgagagcct gacctattgc atctcccgcc gtgcacaggg	9240
tgtcacgttg caagacctgc ctgaaaccga actgcccgct gttctgcagc cggtcgcgga	9300
ggccatggat gcgatcgctg cggccgatct tagccagacg agcgggttcg gcccattcgg	9360
accgcaagga atcggtcaat acactacatg gcgtgatttc atatgcgcga ttgctgatcc	9420
ccatgtgtat cactggcaaa ctgtgatgga cgacaccgtc agtgcgtccg tcgcgcaggc	9480
tctcgatgag ctgatgcttt gggccgagga ctgccccgaa gtccggcacc tcgtgcacgc	9540
ggatttcggc tccaacaatg tcctgacgga caatggccgc ataacagcgg tcattgactg	9600
gagegaggeg atgttegggg atteceaata egaggtegee aacatettet tetggaggee	9660

WO 2004/063359 PCT/EP2004/000099 254/357

gtggttgget tgtatggage agcagacgeg ctacttegag eggaggeate eggagettge	9720
aggategeeg eggeteeggg egtatatget eegeattggt ettgaeeaac tetateagag	9780
cttggttgac ggcaatttcg atgatgcagc ttgggcgcag ggtcgatgcg acgcaatcgt	9840
ccgatccgga gccgggactg tcgggcgtac acaaatcgcc cgcagaagcg cggccgtctg	9900
gaccgatggc tgtgtagaag tactcgccga tagtggaaac cgacgcccca gcactcgtcc	9960
gagggcaaag gaatagagta gatgccgacc gcgggatcga tccacttaac gttactgaaa	10020
tcatcaaaca gcttgacgaa tctggatata agatcgttgg tgtcgatgtc agctccggag	10080
ttgagacaaa tggtgttcag gatctcgata agatacgttc atttgtccaa gcagcaaaga	10140
gtgccttcta gtgatttaat agctccatgt caacaagaat aaaacgcgtt ttcgggttta	10200
cctcttccag atacagctca tctgcaatgc attaatgcat tgactgcaac ctagtaacgc	10260
cttncaggct ccggcgaaga gaagaatagc ttagcagagc tattttcatt ttcgggagac	10320
gagatcaagc agatcaacgg tcgtcaagag acctacgaga ctgaggaatc cgctcttggc	10380
tecaegegae tatatatttg tetetaattg taetttgaea tgeteetett etttaetetg	10440
atagettgae tatgaaaatt eegteaceag encetgggtt egeaaagata attgeatgtt	10500
tetteettga aeteteaage etacaggaca cacatteate gtaggtataa aeetegaaat	10560
canttectae taagatggta tacaatagta accatgcatg gttgcctagt gaatgeteeg	10620
taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttacccaga	10680
atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact	10740
gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt cattttgctt	10800
tgtaaatttc tggtaactgc caccaagaaa tatgaggata ttcgtgatgt tcctcgtggt	10860
agccaaaatg atagcacgtg ataaatgacc accaaatagg acggctaatt gtttgggcac	10920

WO 2004/063359 PCT/EP2004/000099 255/357

aatgaggctg	aacataaccc	cctattggtt	cactatgggg	taaaaaagta	ccaaaataga	10980
ataattgtaa	tgaacttaaa	agcgagggta	gcacccaaaa	gtaagttaga	ttatcacttg	11040
ggatatggag	tatgtattta	gcaaagttat	aaataatagt	caacgcaatt	atttgccccc	11100
aactccagta	acctttcata	aaatgaaaat	accaagcaaa	gaaactttgg	tgtttaccat	11160
tgtgaaaatc	cgggtctatt	gagettgetg	gattgtggtg	gtgtaaccaa	tgtttttca	11220
atagtttttg	atatggtaaa	agaccataaa	gggatagggt	caatgttcca	atcaaatgat	11280
taatcttggt	gttttgggga	aatactacgc	catgcatggc	atcatgagat	gtaataaata	11340
atcccgtata	taaaaatgtt	tgccatagta	taacaggcaa	taacatccaa	aattttagct	11400
ttgagatgtc	aagggaaagt	aataaactca	ggctaatgac	ccatgcgcta	acaatgacaa	11460
tagcaatgaa	aagcccctta	aactgagatt	tacttctcag	tactggagtc	agttttgctt	11520
gatgactgag	tggttgttct	aactggatca	tttctaaaga	gaaggtggaa	caatgttagc	11580
ataattgtgc	ttgagtgagg	actttgaggg	taggtacata	cttgataaag	ttaatgatta	11640
aagagaaaaa	aaaagttttg	gttcaaagca	gaaattgttt	tttaaatcga	ttggtgagaa	11700
aattttttc	tgtttccgca	tcaccaaagc	cacctcagga	atggtcacaa	attattggtc	11760
tgattggacc	ataagcatac	aaaaagttca	ttgaagtata	cttagtggct	tattagactt	11820
ttatcgtttt	ctaacgcgaa	tcagcaatgt	ttcttgtttg	atttactgct	tgctttagat	11880
catttttgtc	tgaaatatta	tgcatttgtt	caaagcggcc	tttgtttcct	ttctttcatg	11940
cttaaacacg	ttgtttattc	catatattac	tttgaatatg	catcaccgca	aagcggaagt	12000
gcaaaataac	aaagaacctc	tttgggttac	acgatcaact	gctattgtga	aaaaaatttc	12060
tttttgaaaa	tttttggaat	aatatctctt	gcaaaaaaga	aattttgtat	atttagtagc	12120
atcaagaaca	. aatgaaagaa	gtgtgggata	acaagaatac	atcatcttta	gacaaaagta	12180

WO 2004/063359 PCT/EP2004/000099 256/357

cgagaaaaat	ctaataagtt	gttatagagg	tctttgtttt	ctttgtgttt	atagacagtt	12240
atttagagtt	tgaaaagtgt	ctctaatgtg	tctttttta	ttattattat	ttcaaatgtt	12300
atgtaatata	gctaaagcta	tagatttgac	attttttcta	aatataaaat	ttcagtcaac	12360
agaaataaat	gacacgagtt	ctttttctct	ctctcaatcc	tgttgatcat	caatctttga	12420
tgtcgtttta	aaacaaatga	atggcattta	gttccttagg	tgtcactcac	atcttgttga	12480
ccagaaaatc	cttattcgcc	ctcaaatctg	ctttattcct	ttcatttgat	ttgatgttta	12540
agtaatgcaa	gcaaacaaaa	aagaaacctt	tcttgcaaag	acaaaagaat	tgttttcaga	12600
ggaaagcaac	tcgttgtcat	tttttaagga	tttagactta	taatcgacac	catagtttgt	12660
ccgttacatt	ttttattgtc	gttttctgat	ttccttttaa	tctttaagca	aaatcaatat	12720
taacttatct	tgtcttccaa	taaaaaatgg	ataccaataa	caataaatcc	ttcacaaaga	12780
aaaaaaaaa	aaactcgaaa	aaagcttggc	gtaatcatgg	tcatagctgt	ttcctgtgtg	12840
aaattgttat	ccgctcacaa	ttccacacaa	catacgagcc	ggaagcataa	agtgtaaagc	12900
ctggggtgcc	taatgagtga	gctaactcac	attaattgcg	ttgcgctcac	tgcccgcttt	12960
ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	ggccaacgcg	cggggagagg	13020
cggtttgcgt	attgggccaa	agacaaaagg	gcgacattca	accgattgag	ggagggaagg	13080
taaatattga	cggaaattat	tcattaaagg	tgaattatca	ccgtcaccga	cttgagccat	13140
ttgggaatta	gagccagcaa	aatcaccagt	agcaccatta	ccattagcaa	ggccggaaac	13200
gtcaccaatg	aaaccatcga	tagcagcacc	gtaatcagta	gcgacagaat	caagtttgcc	13260
tttagcgtca	gactgtagcg	cgttttcatc	ggcattttcg	gtcatagccc	ccttattagc	13320
gtttgccatc	ttttcataat	caaaatcacc	ggaaccagag	ccaccaccgg	aaccgcctcc	13380
ctcagagccg	ccaccctcag	aaccgccacc	: ctcagagcca	ccaccctcag	agccgccacc	13440

WO 2004/063359 PCT/EP2004/000099 257/357

agaaccacca	ccagagccgc	cgccagcatt	gacaggaggc	ccgatctagt	aacatagatg	13500
acaccgcgcg	cgataattta	tcctagtttg	cgcgctatat	tttgttttct	atcgcgtatt	13560
aaatgtataa	ttgcgggact	ctaatcataa	aaacccatct	cataaataac	gtcatgcatt	13620
acatgttaat	tattacatgc	ttaacgtaat	tcaacagaaa	ttatatgata	atcatcgcaa	13680
gaccggcaac	aggattcaat	cttaagaaac	tttattgcca	aatgtttgaa	cgatcgggga	13740
tcatccgggt	ctgtggcggg	aactccacga	aaatatccga	acgcagcaag	atatcgcggt	13800
gcatctcggt	cttgcctggg	cagtcgccgc	cgacgccgtt	gatgtggacg	ccgggcccga	13860
tcatattgtc	gctcaggatc	gtggcgttgt	gcttgtcggc	cgttgctgtc	gtaatgatat	13920
cggcaccttc	gaccgcctgt	tccgcagaga	tecegtggge	gaagaactcc	agcatgagat	13980
ccccgcgctg	gaggatcatc	cagccggcgt	cccggaaaac	gattccgaag	cccaaccttt	14040
catagaaggc	ggcggtggaa	tcgaaatctc	gtgatggcag	gttgggcgtc	gcttggtcgg	14100
tcatttcgaa	ccccagagtc	ccgctcagaa	gaactcgtca	agaaggcgat	agaaggcgat	14160
gegetgegaa	tcgggagcgg	cgataccgta	aagcacgagg	aagcggtcag	cccattcgcc	14220
gccaagctct	tcagcaatat	cacgggtagc	caacgctatg	tcctgatagc	ggtccgccac	14280
acccagccgg	ccacagtcga	tgaatccaga	aaagcggcca	ttttccacca	tgatattcgg	14340
caagcaggca	tegecatggg	tcacgacgag	atcatcgccg	tegggcatge	gcgccttgag	14400
cctggcgaac	agttcggctg	gcgcgagccc	ctgatgctct	tcgtccagat	catcctgatc	14460
gacaagaccg	gcttccatcc	gagtacgtgc	tcgctcgatg	cgatgtttcg	cttggtggtc	14520
gaatgggcag	gtagccggat	caagcgtatg	cagccgccgc	attgcatcag	ccatgatgga	14580
tactttctcg	gcaggagcaa	ggtgagatga	caggagatcc	tgccccggca	cttcgcccaa	14640
tagcagccag	tecetteceg	cttcagtgac	aacgtcgagc	acagetgege	aaggaacgcc	14700

WO 2004/063359 PCT/EP2004/000099 258/357

cgtcgtggcc	agccacgata	gccgcgctgc	ctcgtcctgc	agttcattca	gggcaccgga	14760
caggteggte	ttgacaaaaa	gaaccgggcg	cccctgcgct	gacagccgga	acacggcggc	14820
atcagagcag	ccgattgtct	gttgtgccca	gtcatagccg	aatagcctct	ccacccaagc	14880
ggccggagaa	cctgcgtgca	atccatcttg	ttcaatcatg	cgaaacgatc	cagatccggt	14940
gcagattatt	tggattgaga	gtgaatatga	gactctaatt	ggataccgag	gggaatttat	15000
ggaacgtcag	tggagcattt	ttgacaagaa	atatttgcta	gctgatagtg	accttaggcg	15060
acttttgaac	gcgcaataat	ggtttctgac	gtatgtgctt	agctcattaa	actccagaaa	15120
cccgcggctg	agtggctcct	tcaacgttgc	ggttctgtca	gttccaaacg	taaaacggct	15180
tgtcccgcgt	categgeggg	ggtcataacg	tgactccctt	aattctccgc	tcatgatcag	15240
attgtcgttt	cccgccttca	gtttaaacta	tcagtgtttg	acaggatata	ttggcgggta	15300
aacctaagag	aaaagagcgt	ttattagaat	aatcggatat	ttaaaagggc	gtgaaaaggt	15360
ttatccgttc	gtccatttgt	atgtgcatgc	caaccacagg	gttccccaga	tetggegeeg	15420
gccagcgaga	cgagcaagat	tggccgccgc	ccgaaacgat	ccgacagcgc	gcccagcaca	15480
ggtgcgcagg	caaattgcac	caacgcatac	agcgccagca	gaatgccata	gtgggcggtg	15540
acgtcgttcg	agtgaaccag	atcgcgcagg	aggcccggca	gcaccggcat	aatcaggccg	15600
atgccgacag	cgtcgagcgc	gacagtgctc	agaattacga	tcaggggtat	gttgggtttc	15660
acgtctggcc	tccggaccag	cctccgctgg	tccgattgaa	cgcgcggatt	ctttatcact	15720
gataagttgg	tggacatatt	atgtttatca	gtgataaagt	gtcaagcatg	acaaagttgc	15780
agccgaatac	agtgatccgt	gccgccctgg	acctgttgaa	cgaggtcggc	gtagacggtc	15840
tgacgacacg	caaactggcg	gaacggttgg	gggttcagca	gccggcgctt	tactggcact	15900
tcaggaacaa	gegggegetg	ctcgacgcac	tggccgaagc	catgctggcg	gagaatcata	15960

WO 2004/063359 PCT/EP2004/000099 259/357

cgcattcggt	gccgagagcc	gacgacgact	ggcgctcatt	tctgatcggg	aatgcccgca	16020
gcttcaggca	ggcgctgctc	gcctaccgcg	atggcgcgcg	catccatgcc	ggcacgcgac	16080
cgggcgcacc	gcagatggaa	acggccgacg	cgcagcttcg	cttcctctgc	gaggcgggtt	16140
tttcggccgg	ggacgccgtc	aatgcgctga	tgacaatcag	ctacttcact	gttggggccg	16200
tgcttgagga	gcaggccggc	gacagcgatg	ccggcgagcg	cggcggcacc	gttgaacagg	16260
ctccgctctc	gccgctgttg	cgggccgcga	tagacgcctt	cgacgaagcc	ggtccggacg	16320
cagcgttcga	gcagggactc	gcggtgattg	tcgatggatt	ggcgaaaagg	aggctcgttg	16380
tcaggaacgt	tgaaggaccg	agaaagggtg	acgattgatc	aggaccgctg	ccggagcgca	16440
acccactcac	tacagcagag	ccatgtagac	aacatcccct	cccctttcc	accgcgtcag	16500
acgcccgtag	cagcccgcta	cgggcttttt.	catgccctgc	cctagcgtcc	aagcctcacg	16560
gccgcgctcg	gcctctctgg	cggccttctg	gcgctcttcc	gattactaga	tcactgactc	16620
gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct [.]	cactcaaagg	cggtaatacg	16680
gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	tgagcaaaag	gccagcaaaa	16740
ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	cataggctcc	gccccctga	16800
cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	aacccgacag	gactataaag	16860
ataccaggcg	tttccccctg	gaageteeet	cgtgcgctct	cctgttccga	cectgeeget	16920
taccggatac	ctgtccgcct	ttctcccttc	gggaagcgtg	gcgcttttcc	gctgcataac	16980
cctgcttcgg	ggtcattata	gcgattttt	cggtatatcc	atcctttttc	gcacgatata	17040
caggattttg	ccaaagggtt	cgtgtagact	ttccttggtg	tatccaacgg	cgtcagccgg	17100
gcaggatagg	tgaagtaggc	ccacccgcga	gcgggtgttc	cttcttcact	gtcccttatt	17160
cgcacctggc	ggtgctcaac	gggaatcctg	ctctgcgagg	ctggccggct	accgccggcg	17220

taacagatga gggcaagegg atggctgatg aaaccaagec aaccaggaag ggcageccac 17280
ctatcaaggt gtactgeett eeagacgaac gaagagegat tgaggaaaaag geggeggegg 17340
ceggcatgag eetgteggee taeetgetgg eegteggeea gggetacaaa ateaegggeg 17400
tegtggacta tgagcaegte egegagetgg eeegeateaa tggegaeetg ggeegeetgg 17460
geggeetget gaaactetgg etcacegaeg aeeeggeae ggegggtte ggtgatgeea 17520
egateetege eetgetggeg aagategaag agaageagga egagettgge aaggteatga 17580
tgggegtggt eegeeegggg geagageeat gaettttta geegetaaaa eggeeggggg 17640
gtgegegtga ttgeeaagea egteeeeatg egeteeatea agaagagega ettegegaag 17700
ctggtgaagt acateaecga egageaagge aagacegage geetttgega egetea 17756

```
<210> 48
```

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<211> 17118

<212> DNA

<223> n is a, c, g, or t

<400> 48						
ccgggctggt	tgccctcgcc	gctgggctgg	cggccgtcta	tggccctgca	aacgcgccag	60
aaacgccgtc	gaagccgtgt	gcgagacacc	geggeegeeg	gcgttgtgga	tacctcgcgg	120
aaaacttggc	cctcactgac	agatgagggg	cggacgttga	cacttgaggg	gccgactcac	180
ceggegegge	gttgacagat	gaggggcagg	ctcgatttcg	geeggegaeg	tggagctggc	240
cagcctcgca	aatcggcgaa	aacgcctgat	tttacgcgag	tttcccacag	atgatgtgga	300
caagcctggg	gataagtgcc	ctgcggtatt	gacacttgag	gggcgcgact	actgacagat	360
gaggggcgcg	atccttgaca	cttgaggggc	agagtgctga	cagatgaggg	gcgcacctat	420
tgacatttga	ggggctgtcc	acaggcagaa	aatccagcat	ttgcaagggt	ttccgcccgt	480
ttttcggcca	ccgctaacct	gtcttttaac	ctgcttttaa	accaatattt	ataaaccttg	540
tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcgcac	gccgaagggg	ggtgccccc	600
cttctcgaac	cctcccggcc	cgctaacgcg	ggcctcccat	cccccaggg	gctgcgcccc	660
tcggccgcga	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	cattgccggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	cggcggcctg	840
ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaatttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg	gacctttaca	gaattactct	atgaagcgcc	atatttaaaa	agctaccaag	1080
acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	aatactgata	1140
agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	ggggcaaggc	1200

WO 2004/063359 PCT/EP2004/000099 262/357

	1060
ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga	1260
ctaatgcttg aaacccagga caataacctt atagcttgta aattctatca taattgggta	1320
atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac tttgtcatgc	1380
agetecaceg attttgagaa egacagegae tteegteeea geegtgeeag gtgetgeete	1440
agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac gtgcagcttt	1500
cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac cacgtcaaag	1560
ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc gaatacgtgc	1620
gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg gcgcgattta	1680
gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc actgcccggc	1740
tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa atcgtgttga	1800
ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg gccatatcaa	1860
tgattttctg gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt tgccatgttt	1920
tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg ttacgcacca	1980
ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact ggagcacctc	2040
aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat tgtggtttca	2100
aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt gaaaaagctg	2160
ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagttc gtcttgttat	2220
aattagcttc ttggggtatc tttaaatact gtagaaaaga ggaaggaaat aataaatggc	2280
taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct gcgtaaaaga	2340
tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg aaaacctata	2400
tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac gggaaaagga	2460

WO 2004/063359 PCT/EP2004/000099 263/357

catgatgcta	tggctggaag	gaaagctgcc	tgttccaaag	gtcctgcact	ttgaacggca	2520
tgatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac	ttactgaata	acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt	aaagatccgc	gcgagctgta	tgatttttta	aagacggaaa	agcccgaaga	2820
ggaacttgtc	ttttcccacg	gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc	tttattgatc	ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc	cggtcgatca	gggaggatat	cggggaagaa	cagtatgtcg	agctattttt	3000
tgacttactg	gggatcaagc	ctgattggga	gaaaataaaa	tattatattt	tactggatga	3060
attgttttag	tacctagatg	tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tcttccgcat	caagtgtttt	ggctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt	attcgtgcag	ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg	gaccgacttc	attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa	tcaggaataa	gggcacattg	ccccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc	cgaggatgcc	gaaaccatcg	caagccgcac	cgtcatgcgt	gcgccccgcg	3480
aaaccttcca	gtccgtcggc	tegatggtee	agcaagctac	ggccaagatc	gagegegaca	3540
gcgtgcaact	ggctcccct	geeetgeeeg	cgccatcggc	cgccgtggag	cgttcgcgtc	3600
gtctcgaaca	ggaggcggca	ggtttggcga	agtcgatgac	catcgacacg	cgaggaacta	3660
tgacgaccaa	gaagcgaaaa	accgccggcg	aggacctggc	aaaacaggtc	agcgaggcca	3720

agcaggccgc	gttgctgaaa	cacacgaagc	agcagatcaa	ggaaatgcag	ctttccttgt	3780
tcgatattgc	gccgtggccg	gacacgatgc	gagcgatgcc	aaacgacacg	gcccgctctg	3840
ccctgttcac	cacgcgcaac	aagaaaatcc	cgcgcgaggc	gctgcaaaac	aaggtcattt	3900
tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctatc	ggcgagccga	4020
tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggc	cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg	gcacctggaa	teggtgtege	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200
gcaagaaaac	gtcccgttgc	caggtcctga	tcgacgagga	aatcgtcgtg	ctgtttgctg	4260
gcgaccacta	cacgaaattc	atatgggaga	agtaccgcaa	gctgtcgccg	acggcccgac	4320
ggatgttcga	ctatttcagc	tegeaceggg	agccgtaccc	gctcaagctg	gaaaccttcc	4380
gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
cctgcgaaga	gttgcgaggc	agcggcctgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	gcagccagcg	4560
ctttactggc	atttcaggaa	caagcgggca	ctgctcgacg	cacttgcttc	gctcagtatc	4620
gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
tgtgattaag	gctcagattc	gacggcttgg	agcggccgac	gtgcaggatt	tccgcgagat	4740
ccgattgtcg	gccctgaaga	. aagctccaga	gatgttcggg	teegtttaeg	agcacgagga	4800
gaaaaagccc	: atggaggcgt	tegetgaaeg	gttgcgagat	gccgtggcat	teggegeeta	4860
· categaegge	gagatcattg	ggetgteggt	: cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggc	g catctgtccg	gegttttegt	ggagcccgaa	cagegagge	gaggggtcgc	4980

WO 2004/063359 PCT/EP2004/000099 265/357

cggtatgctg ctgcgggcgt tgccggcggg tttattgctc gtgatgatcg tccgacagat 5040 tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100 ctggagcttg ttgtttattt cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg 5160 cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220 attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tggtgttgac 5280 accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340 ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400 cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460 5520 gccaatcccg atgcctacag gaaccaatgt teteggeetg gegtggeteg geetgategg 5580 agegggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt ttccttactg ggctttctca gccccagatc tggggtcgat cagccgggga tgcatcaggc 5640 5700 cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5760 tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820 5880 cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940 6000 gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6060 tetgeegeet tacaaegget etecegetga egeegteeeg gaetgatggg etgeetgtat 6120 cgagtggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6180 tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6240 taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc

WO 2004/063359 PCT/EP2004/000099 267/357

7560 gcactattga tcatccgata gctctgcaaa gggcgttgca caatgcaagt caaacgttgc 7620 tagcagttcc aggtggaatg ttatgatgag cattgtatta aatcaggaga tatagcatga tctctagtta gctcaccaca aaagtcagac ggcgtaacca aaagtcacac aacacaagct 7680 gtaaggattt cggcacggct acggaagacg gagaagccac cttcagtgga ctcgagtacc 7740 atttaattct atttgtgttt gatcgagacc taatacagcc cctacaacga ccatcaaagt 7800 cgtatagcta ccagtgagga agtggactca aatcgacttc agcaacatct cctggataaa 7860 7920 ctttaagcct aaactataca gaataagata ggtggagagc ttataccgag ctcccaaatc tgtccagatc atggttgacc ggtgcctgga tcttcctata gaatcatcct tattcgttga 7980 cctagctgat tctggagtga cccagagggt catgacttga gcctaaaatc cgccgcctcc 8040 8100 accatttgta gaaaaatgtg acgaactcgt gagctctgta cagtgaccgg tgactctttc tggcatgcgg agagacggac ggacgcagag agaagggctg agtaataagc cactggccag 8160 acagetetgg eggetetgag gtgeagtgga tgattattaa teegggaeeg geegeeeete 8220 cgccccgaag tggaaaggct ggtgtgcccc tcgttgacca agaatctatt gcatcatcgg 8280 8340 agaatatgga gcttcatcga atcaccggca gtaagcgaag gagaatgtga agccaggggt gtatagccgt cggcgaaata gcatgccatt aacctaggta cagaagtcca attgcttccg 8400 atctggtaaa agattcacga gatagtacct tctccgaagt aggtagagcg agtacccggc 8460 gcgtaagctc cctaattggc ccatccggca tctgtagggc gtccaaatat cgtgcctctc 8520 ctgctttgcc cggtgtatga aaccggaaag gccgctcagg agctggccag cggcgcagac 8580 cgggaacaca agctggcagt cgacccatcc ggtgctctgc actcgacctg ctgaggtccc 8640 8700 tcagtccctg gtaggcagct ttgccccgtc tgtccgcccg gtgtgtcggc ggggttgaca aggtcgttgc gtcagtccaa catttgttgc catattttcc tgctctcccc accagctgct 8760

cttttctttt ctct	ttettt teceatette	agtatattca	tcttcccatc	caagaaçctt	8820
tatttcccct aagt	aagtac tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt tcag	sttegag ettteceaet	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac atca	accatgc ctgaactcac	: cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac agcg	gteteeg acetgatgea	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat gtag	ggagggc gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat cgtt	catgttt atcggcactt	tgcatcggcc	gcgctcccga	ttccggaagt	9180
gcttgacatt gggg	gaattca gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg caag	gacetge etgaaacega	a actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat gcga	ategetg eggeegate	tagccagacg	agcgggttcg	gcccattcgg	9360
accgcaagga atcg	ggtcaat acactacat	g gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat cact	tggcaaa ctgtgatgg	a cgacaccgtc	agtgcgtccg	tegegeagge	9480
tctcgatgag ctga	atgettt gggeegagg	a ctgccccgaa	gtccggcacc	tegtgeaege	9540
ggatttegge tee	aacaatg teetgaegg	a caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg atg	ttcgggg attcccaat	a cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct tgt	atggagc agcagacgc	g ctacttcgag	cggaggcatc	cggagcttgc	9720
aggategeeg egg	ctccggg cgtatatgc	t ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac ggc	aatttcg atgatgcag	c ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga gcc	gggactg tcgggcgta	c acaaatcgcc	cgcagaagcg	gegeegtetg	9900
gaccgatggc tgt	gtagaag tactcgccg	a tagtggaaac	: cgacgcccca	gcactcgtcc	9960
gagggcaaag gaa	utagagta gatgeegae	c gcgggatcga	tccacttaac	gttactgaaa	10020

WO 2004/063359 PCT/EP2004/000099 269/357

tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080
ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttegggttta	10200
cctcttccag	atacagctca	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagagc	tattttcatt	ttcgggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380
tccacgcgac	tatatatttg	tctctaattg	tactttgaca	tgctcctctt	ctttactctg	10440
atagcttgac	tatgaaaatt	ccgtcaccag	cncctgggtt	cgcaaagata	attgcatgtt	10500
tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
canttcctac	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactctc	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagcgc	ccactccaca	tctccactcg	acctgcaggc	atgcaagctt	gagattaaaa	10800
tagataagga	aaagaaagtg	aaaagaaatt	cggaagcatg	gcacattctt	ctttttataa	10860
atacatgcct	gactttcttt	ttccatcgat	atgatatatg	catatgatag	atatacaagc	10920
aatcttcttc	aaggagtttg	aaattttgtc	ctccaggagc	aaaaaaaagt	tttttttat	10980
acatgtttgt	acacaagaat	agttaccaat	ttgctttggt	cttacgtgct	gcaagtttat	11040
atcgttttca	atttctttgt	ctttacattt	tetttgteet	ttatctttcc	tcatttagtc	11100
tttgggagaa	ttaggaaaag	ggagcggaaa	ggtaagaaat	gcttgcgtat	tttactaatt	11160
cggcaaacat	ccaatttggc	aaacagcagc	ctgtgcaacg	ctctcgagat	gacagtatct	11220
ttgattacac	tctaaatctc	gatgacccga	ccaaaaagag	cgaacaaaga	aataatcttg	11280

WO 2004/063359 PCT/EP2004/000099 270/357

tgcattcgaa	tatgatggaa	gattttttcc	cccttattct	aaatgttgac	atagcgtgta	11340
tgttatataa	acaaaaagaa	attgtacaaa	ctttctttc	ttctctttt	attttatctc	11400
tatgatccag	ttagaacaac	cactcagtca	tcaagcaaaa	ctgactccag	tactgagaag	11460
taaatctcag	tttaaggggc	ttttcattgc	tattgtcatt	gttagcgcat	gggtcattag	11520
cctgagttta	ttactttccc	ttgacatctc	aaagctaaaa	ttttggatgt	tattgcctgt	11580
tatactatgg	caaacatttt	tatatacggg	attatttatt	acatctcatg	atgccatgca	11640
tggcgtagta	tttccccaaa	acaccaagat	taatcatttg	attggaacat	tgaccctatc	11700
cctttatggt	cttttaccat	atcaaaaact	attgaaaaaa	cattggttac	accaccacaa	11760
tccagcaagc	tcaatagacc	cggattttca	caatggtaaa	caccaaagtt	tctttgcttg	11820
gtattttcat	tttatgaaag	gttactggag	ttgggggcaa	ataattgcgt	tgactattat	11880
ttataacttt	gctaaataca	tactccatat	cccaagtgat	aatctaactt	acttttgggt	11940 ·
gctaccctcg	cttttaagtt	cattacaatt	attctatttt	ggtacttttt	taccccatag	12000
tgaaccaata	gggggttatg	ttcagcctca	ttgtgcccaa	acaattagcc	gtcctatttg	12060
gtggtcattt	atcacgtgct	atcattttgg	ctaccacgag	gaacatcacg	aatatcctca	12120
tatttcttgg	tggcagttac	cagaaattta	caaagcaaaa	tagaagcttg	gcgtaatcat	12180
ggtcatagct	gtttcctgtg	tgaaattgtt	atccgctcac	aattccacac	aacatacgag	12240
ccggaagcat	aaagtgtaaa	gcctggggtg	cctaatgagt	gagctaactc	acattaattg	12300
cgttgcgctc	actgcccgct	ttccagtcgg	gaaacctgto	gtgccagctg	cattaatgaa	12360
tcggccaacg	cgcggggaga	ggcggtttgc	gtattgggco	aaagacaaaa	gggcgacatt	12420
caaccgattg	agggagggaa	ggtaaatatt	gacggaaatt	attcattaaa	ggtgaattat	12480
caccgtcacc	gacttgagcc	atttgggaat	tagagccago	: aaaațcacca	gtagcaccat	12540

WO 2004/063359 PCT/EP2004/000099 271/357

taccattagc aaggccggaa acgtcaccaa tgaaaccatc gatagcagca ccgtaatcag 12600 tagcgacaga atcaagtttg cctttagcgt cagactgtag cgcgttttca tcggcatttt 12660 cggtcatagc ccccttatta gcgtttgcca tcttttcata atcaaaatca ccggaaccag 12720 agceaccace ggaaccgcct cecteagage egecaccete agaaccgcca cecteagage 12780 caccaccete agageegeca ceagaaceae caccagagee geegecagea ttgacaggag 12840 gcccgatcta gtaacataga tgacaccgcg cgcgataatt tatcctagtt tgcgcgctat 12900 attttgtttt ctatcgcgta ttaaatgtat aattgcggga ctctaatcat aaaaacccat 12960 ctcataaata acgtcatgca ttacatgtta attattacat gcttaacgta attcaacaga 13020 aattatatga taatcatcgc aagaccggca acaggattca atcttaagaa actttattgc 13080 caaatgtttg aacgatcggg gatcatccgg gtctgtggcg ggaactccac gaaaatatcc 13140 gaacgcagca agatatcgcg gtgcatctcg gtcttgcctg ggcagtcgcc gccgacgccg 13200 ttgatgtgga cgccgggccc gatcatattg tcgctcagga tcgtggcgtt gtgcttgtcg 13260 . gccgttgctg tcgtaatgat atcggcacct tcgaccgcct gttccgcaga gatcccgtgg 13320 gcgaagaact ccagcatgag atccccgcgc tggaggatca tccagccggc gtcccggaaa 13380 acgattccga agcccaacct ttcatagaag gcggcggtgg aatcgaaatc tcgtgatggc 13440 aggttgggcg tcgcttggtc ggtcatttcg aaccccagag tcccgctcag aagaactcgt 13500 caagaaggcg atagaaggcg atgcgctgcg aatcgggagc ggcgataccg taaagcacga 13560 ggaagcggtc agcccattcg ccgccaagct cttcagcaat atcacgggta gccaacgcta 13620 tgtcctgata gcggtccgcc acacccagcc ggccacagtc gatgaatcca gaaaagcggc 13680 cattttccac catgatattc ggcaagcagg catcgccatg ggtcacgacg agatcatcgc 13740 cgtcgggcat gcgcgccttg agcctggcga acagttcggc tggcgcgagc ccctgatgct 13800

WO 2004/063359 PCT/EP2004/000099 272/357

cttcgtccag	atcatcctga	tcgacaagac	cggcttccat	ccgagtacgt	gctcgctcga	13860
tgcgatgttt	cgcttggtgg	tcgaatgggc	aggtagccgg	atcaagcgta	tgcagccgcc	13920
gcattgcatc	agccatgatg	gatactttct	cggcaggagc	aaggtgagat	gacaggagat	13980
cctgccccgg	cacttcgccc	aatagcagcc	agtecettee	cgcttcagtg	acaacgtcga	14040
gcacagctgc	gcaaggaacg	cccgtcgtgg	ccagccacga	tagccgcgct	gcctcgtcct	14100
gcagttcatt	cagggcaccg	gacaggtcgg	tettgacaaa	aagaaccggg	cgcccctgcg	14160
ctgacagccg	gaacacggcg	gcatcagagc	agccgattgt	ctgttgtgcc	cagtcatagc	14220
cgaatagcct	ctccacccaa	geggeeggag	aacctgcgtg	caatccatct	tgttcaatca	14280
tgcgaaacga	tccagatccg	gtgcagatta	tttggattga	gagtgaatat	gagactctaa	14340
ttggataccg	aggggaattt	atggaacgtc	agtggagcat	ttttgacaag	aaatatttgc	14400
tagctgatag	tgaccttagg	cgacttttga	acgcgcaata	atggtttctg	acgtatgtgc	14460
ttagctcatt	. aaactccaga	aacccgcggc	tgagtggctc	cttcaacgtt	geggttetgt	14520
cagttccaaa	cgtaaaacgg	cttgtcccgc	gtcatcggcg	ggggtcataa	cgtgactccc	14580
ttaattctcc	gctcatgatc	agattgtcgt	ttcccgcctt	cagtttaaac	tatcagtgtt	14640
tgacaggata	tattggcggg	taaacctaag	agaaaagagc	gtttattaga	ataatcggat	14700
atttaaaagg	g gcgtgaaaag	gtttatccgt	tcgtccattt	gtatgtgcat	gccaaccaca	14760
gggttcccca	a gatctggcgc	cggccagcga	gacgagcaag	attggccgcc	gcccgaaacg	14820
atccgacage	gegeeeagea	caggtgcgca	ggcaaattgc	: accaacgcat	acagegeeag	14880
cagaatgcc	a tagtgggcgg	tgacgtcgtt	cgagtgaaco	: agatcgcgca	a ggaggcccgg	14940
cagcaccgg	c ataatcaggo	e cgatgccgac	agcgtcgago	gegaeagtge	c tcagaattac	15000
gatcagggg	t atgttgggtt	: tcacgtctgg	cctccggaco	agectecge	t ggtccgattg	15060

WO 2004/063359 PCT/EP2004/000099 273/357

aacgcgcgga	ttctttatca	ctgataagtt	ggtggacata	ttatgtttat	cagtgataaa	15120
gtgtcaagca	tgacaaagtt	gcagccgaat	acagtgatcc	gtgccgccct	ggacctgttg	15180
aacgaggtcg	gcgtagacgg	tctgacgaca	cgcaaactgg	cggaacggtt	gggggttcag	15240
cagccggcgc	tttactggca	cttcaggaac	aagcgggcgc	tgctcgacgc	actggccgaa	15300
gccatgctgg	cggagaatca	tacgcattcg	gtgccgagag	ccgacgacga	ctggcgctca	15360
tttctgatcg	ggaatgcccg	cagcttcagg	caggcgctgc	tegeetaceg	cgatggcgcg	15420
cgcatccatg	ccggcacgcg	accgggcgca	ccgcagatgg	aaacggccga	cgcgcagctt	15480
cgcttcctct	gcgaggcggg	tttttcggcc	ggggacgccg	tcaatgcgct	gatgacaatc	15540
agctacttca	ctgttggggc	cgtgcttgag	gagcaggccg	gcgacagcga	tgccggcgag	15600
cgcggcggca	ccgttgaaca	ggctccgctc	tcgccgctgt	tgcgggccgc	gatagacgcc	15660
ttcgacgaag	ccggtccgga	cgcagcgttc	gagcagggac	tegeggtgat	, tgtcgatgga	15720
ttggcgaaaa	ggaggctcgt	tgtcaggaac	gttgaaggac	cgagaaaggg	tgacgattga	15780
tcaggaccgc	tgccggagcg	caacccactc	actacagcag	agccatgtag	acaacatccc	15840
ctccccttt	ccaccgcgtc	agacgcccgt	agcagcccgc	tacgggcttt	ttcatgccct	15900
gccctagcgt	ccaagcctca	cggccgcgct	cggcctctct	ggcggccttc	tggcgctctt	15960
ccgcttcctc	gctcactgac	tegetgeget	cggtcgttcg	gctgcggcga	gcggtatcag	16020
ctcactcaaa	ggcggtaata	cggttatcca	cagaatcagg	ggataacgca	ggaaagaaca	16080
tgtgagcaaa	aggccagcaa	aaggccagga	accgtaaaaa	ggccgcgttg	ctggcgtttt	16140
tccataggct	ccgcccccct	gacgagcatc	acaaaaatcg	acgctcaagt	cagaggtggc	16200
gaaacccgac	aggactataa	agataccagg	cgtttccccc	tggaagctcc	ctcgtgcgct	16260
ctcctgttcc	gaccctgccg	cttaccggat	acctgtccgc	ctttctccct	tcgggaagcg	16320

tggcgctttt	ccgctgcata	accctgcttc	ggggtcatta	tagcgatttt	ttcggtatat	16380
ccatcctttt	tcgcacgata	tacaggattt	tgccaaaggg	ttcgtgtaga	ctttccttgg	16440
tgtatccaac	ggcgtcagcc	gggcaggata	ggtgaagtag	geccaecege	gagcgggtgt	16500
tccttcttca	ctgtccctta	ttcgcacctg	gcggtgctca	acgggaatcc	tgctctgcga	16560
ggctggccgg	ctaccgccgg	cgtaacagat	gagggcaagc	ggatggctga	tgaaaccaag	16620
ccaaccagga	agggcagccc	acctatcaag	gtgtactgcc	ttccagacga	acgaagagcg	16680
attgaggaaa	aggcggcggc	ggccggcatg	agcctgtcgg	cctacctgct	ggccgtcggc	16740
cagggctaca	aaatcacggg	cgtcgtggac	tatgagcacg	tccgcgagct	ggcccgcatc	16800
aatggcgacc	tgggccgcct	gggcggcctg	ctgaaactct	ggctcaccga	cgacccgcgc	16860
acggcgcggt	tcggtgatgc	cacgatcctc	gccctgctgg	cgaagatcga	agagaagcag	16920
gacgagcttg	gcaaggtcat	gatgggcgtg	gtccgcccga	gggcagagcc	atgacttttt	16980
tagccgctaa	aacggccggg	gggtgcgcgt	gattgccaag	cacgtcccca	tgcgctccat	17040
caagaagagc	gacttcgcgg	agctggtgaa	gtacatcacc	gacgagcaag	gcaagaccga	17100
gcgcctttgc	gacgctca					17118

<210> 49

<211> 18449

<212> DNA

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 49

60 gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct 180 cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240 tegggaagee agegactggt gaceteeage gteeetttaa ggetgeeaac agetttetea 300 360 qccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 420 480 atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540 600 aatattctca agtatctcga gtttgaactt attccctgtg aaccttttat tcaccaatga gcattggaat gaacatgaat ctgaggactg caatcgccat gaggttttcg aaatacatcc 660 720 ggatgtcgaa ggcttggggc acctgcgttg gttgaattta gaacgtggca ctattgatca 780 tccgataget ctgcaaaggg cgttgcacaa tgcaagtcaa acgttgctag cagttccagg tggaatgtta tgatgagcat tgtattaaat caggagatat agcatgatct ctagttagct 840 caccacaaaa gtcagacggc gtaaccaaaa gtcacacaac acaagctgta aggatttcgg 900

WO 2004/063359 PCT/EP2004/000099 276/357

				•		
cacggctacg	gaagacggag	aagccacctt	cagtggactc	gagtaccatt	taattctatt	960
tgtgtttgat	cgagacctaa	tacagcccct	acaacgacca	tcaaagtcgt	atagetacea	1020
gtgaggaagt	ggactcaaat	cgacttcagc	aacateteet	ggataaactt	taagcctaaa	1080
ctatacagaa	taagataggt	ggagagctta	taccgagctc	ccaaatctgt	ccagatcatg	1140
gttgaccggt	geetggatet	tcctatagaa	tcatccttat	tcgttgacct	agctgattct	1200
ggagtgaccc	agagggtcat	gacttgagcc	taaaatccgc	cgcctccacc	atttgtagaa	1260
aaatgtgacg	aactcgtgag	ctctgtacag	tgaccggtga	ctctttctgg	catgcggaga	1320
gacggacgga	cgcagagaga	agggctgagt	aataagccac	tggccagaca	gctctggcgg	1380
ctctgaggtg	cagtggatga	ttattaatcc	gggaccggcc	gcccctccgc	cccgaagtgg	1440
aaaggctggt	gtgcccctcg	ttgaccaaga	atctattgca	tcatcggaga	atatggagct	1500
tcatcgaatc	accggcagta	agcgaaggag	aatgtgaagc	caggggtgta	tagccgtcgg	1560
cgaaatagca	tgccattaac	ctaggtacag	aagtccaatt	gcttccgatc	tggtaaaaga	1620
ttcacgagat	agtaccttct	ccgaagtagg	tagagcgagt	acccggcgcg	taagctccct	1680
aattggccca	tccggcatct	gtagggcgtc	caaatategt	gectetectg	ctttgcccgg	1740
tgtatgaaac	cggaaaggcc	gctcaggagc	tggccagcgg	cgcagaccgg	gaacacaagc	1800
tggcagtcga	cccatccggt	gctctgcact	cgacctgctg	aggtccctca	gtccctggta	1860
ggcagctttg	ccccgtctgt	ccgcccggtg	tgtcggcggg ·	gttgacaagg	tegttgegte	1920
agtccaacat	ttgttgccat	attttcctgc	tctccccacc	agctgctctt	ttcttttctc	1980
tttcttttcc	catcttcagt	atattcatct	teccatecaa	gaacctttat	ttcccctaag	2040
taagtacttt	gctacatcca	tactccatcc	ttcccatccc	ttattccttt	gaacctttca	2100
gttcgagctt	tcccacttca	tcgcagcttg	actaacagct	accccgcttg	agcagacatc	2160

WO 2004/063359 PCT/EP2004/000099 277/357

accatgectg aacteacege gacgtetgte gagaagttte tgategaaaa gttegacage	2220
gtctccgacc tgatgcagct ctcggagggc gaagaatctc gtgctttcag cttcgatgta	2280
ggagggcgtg gatatgtcct gcgggtaaat agctgcgccg atggtttcta caaagatcgt	2340
tatgtttatc ggcactttgc atcggccgcg ctcccgattc cggaagtgct tgacattggg	2400
gaattcagcg agagcctgac ctattgcatc tecegecgtg cacagggtgt cacgttgcaa	2460
gacctgcctg aaaccgaact gcccgctgtt ctgcagccgg tcgcggaggc catggatgcg	2520
atcgctgcgg ccgatcttag ccagacgagc gggttcggcc cattcggacc gcaaggaatc	2580
ggtcaataca ctacatggcg tgatttcata tgcgcgattg ctgatcccca tgtgtatcac	2640
tggcaaactg tgatggacga caccgtcagt gcgtccgtcg cgcaggctct cgatgagctg	2700
atgetttggg cegaggaetg eecegaagte eggeaeeteg tgeaegegga ttteggetee	2760
aacaatgtcc tgacggacaa tggccgcata acagcggtca ttgactggag cgaggcgatg	2820
ttcggggatt cccaatacga ggtcgccaac atcttcttct ggaggccgtg gttggcttgt	2880
atggagcagc agacgcgcta cttcgagcgg aggcatccgg agcttgcagg atcgccgcgg	2940
ctccgggcgt atatgctccg cattggtctt gaccaactct atcagagctt ggttgacggc	3000
aatttcgatg atgcagcttg ggcgcagggt cgatgcgacg caatcgtccg atccggagcc	3060
gggactgtcg ggcgtacaca aatcgcccgc agaagcgcgg ccgtctggac cgatggctgt	3120
gtagaagtac tcgccgatag tggaaaccga cgccccagca ctcgtccgag ggcaaaggaa	3180
tagagtagat gccgaccgcg ggatcgatcc acttaacgtt actgaaatca tcaaacagct	3240
tgacgaatct ggatataaga tcgttggtgt cgatgtcagc tccggagttg agacaaatgg	3300
tgttcaggat ctcgataaga tacgttcatt tgtccaagca gcaaagagtg ccttctagtg	3360
atttaatagc tccatgtcaa caagaataaa acgcgttttc gggtttacct cttccagata	3420

WO 2004/063359 PCT/EP2004/000099 278/357

cagctcatct	gcaatgcatt	aatgcattga	ctgcaaccta	gtaacgcctt	ncaggctccg	3480
gcgaagagaa	gaatagctta	gcagagctat	tttcattttc	gggagacgag	atcaagcaga	3540
tcaacggtcg	tcaagagacc	tacgagactg	aggaatccgc	tcttggctcc	acgcgactat	3600
atatttgtct	ctaattgtac	tttgacatgc	tectettett	tactctgata	gcttgactat	3660
gaaaattccg	tcaccagcnc	ctgggttcgc	aaagataatt	gcatgtttct	tccttgaact	3720
ctcaagccta	caggacacac	attcatcgta	ggtataaacc	tcgaaatcan	ttcctactaa	3780
gatggtatac	aatagtaacc	atgcatggtt	gcctagtgaa	tgctccgtaa	cacccaatac	3840
gccggccgaa	acttttttac	aactctccta	tgagtcgttt	acccagaatg	cacaggtaca	3900
cttgtttaga	ggtaatcctt	ctttctagct	agaagtcctc	gtgtactgtg	taagcgccca	3960
ctccacatct	ccactcgacc	tgcaggcatg	caaagcttga	gattaaaata	gataaggaaa	4020
agaaagtgaa	aagaaattcg	gaagcatggc	acattcttct	ttttataaat	acatgcctga	4080
ctttctttt	ccatcgatat	gatatatgca	tatgatagat	atacaagcaa	tcttcttcaa	4140
ggagtttgaa	attttgtcct	ccaggagcaa	aaaaaagttt	ttttttatac	atgtttgtac	4200
acaagaatag	ttaccaattt	gctttggtct	tacgtgctgc	aagtttatat	cgttttcaat	4260
ttctttgtct	ttacattttc	tttgtccttt	atctttcctc	atttagtctt	tgggagaatt	4320
aggaaaaggg	agcggaaagg	taagaaatgc	ttgcgtattt	tactaattcg	gcaaacatcc	4380
aatttggcaa	acagcagcct	gtgcaacgct	ctcgagatga	cagtatcttt	gattacactc	4440
taaatctcga	tgacccgacc	aaaaagagcg	aacaaagaaa	taatcttgtg	cattcgaata	4500
tgatggaaga	ttttttcccc	cttattctaa	atgttgacat	agcgtgtatg	ttatataaac	4560
aaaaagaaat	tgtacaaact	ttettttett	ctctttttat	tttatctcta	tgctgtcgaa	4620
gctgcagtca	atcagcgtca	aggecegeeg	cgttgaacta	gcccgcgaca	tcacgcggcc	4680

WO 2004/063359 PCT/EP2004/000099 279/357

caaagtctgc ctgcatgctc agcggtgctc gttagttcgg ctgcgagtgg cagcaccaca 4740 gacagaggag gcgctgggaa ccgtgcaggc tgccggcgcg ggcgatgagc acagcgccga 4800 tgtagcactc cagcagcttg accgggctat cgcagagcgt cgtgcccggc gcaaacggga 4860 gcagctgtca taccaggctg ccgccattgc agcatcaatt ggcgtgtcag gcattgccat 4920 4980 cttcgccacc tacctgagat ttgccatgca catgaccgtg ggcggcgcag tgccatgggg 5040 tgaagtggct ggcactctcc tcttggtggt tggtggcgcg ctcggcatgg agatgtatgc 5100 ccgctatgca cacaaagcca tctggcatga gtcgcctctg ggctggctgc tgcacaagag 5160 ccaccacaca cctcgcactg gaccctttga agccaacgac ttgtttgcaa tcatcaatgg 5220 actgcccgcc atgctcctgt gtacctttgg cttctggctg cccaacgtcc tgggggcggc ctgctttgga gcggggctgg gcatcacgct atacggcatg gcatatatgt ttgtacacga 5280 tggcctggtg cacaggcgct ttcccaccgg gcccatcgct ggcctgccct acatgaagcg 5340 cctgacagtg gcccaccagc tacaccacag cggcaagtac ggtggcgcgc cctggggtat 5400 5460 gttcttgggt ccacaggagc tgcagcacat tccaggtgcg gcggaggagg tggagcgact ggtcctggaa ctggactggt ccaagcgggc gattgtgact gatagcgaga ctctgggtcg 5520 5580 atgttatctg cctcaacaat ggcttagaaa agaagaaaca gaacaaatac agcaaggcaa 5640 cgcccgtagc ctaggtgatc aaagactgtt gggcttgtct ctgaagcttg taggaaaggc agacgctatc atggtgagag ctaagaaggg cattgacaag ttgccggcaa actgtcaagg 5700 5760 cggtgtacga gctgcttgcc aagtatatgc tgcaattgga tctgtactca agcagcagaa gacaacatat cctacaagag ctcatctaaa aggaagcgaa cgtgccaaga ttgctctgtt 5820 gagtgtatac aacctctatc aatctgaaga caagcctgtg gctctccgtc aagctagaaa 5880 gattaagagt ttttttgttg attagtgaat ttttgtttta tttatgtctg atagttcaat 5940

WO 2004/063359 PCT/EP2004/000099 280/357

aaagagacaa	cacatacaat	ataaaatcat	tgtctttaaa	tgttaattta	gtagagtgta	6000
aagcctgcat	tttttttgta	cgcataaaca	atgaattcac	cccgcttctg	gtttttaaat	6060
aattatgtca	aactagggaa	aattctttt	tttctcttcg	ttctttttt	ggcttgttgt	6120
ggagtcacag	gcttgtcttc	agattgatag	aggttgtata	cactcaacag	agcaatcttg	6180
gcacgttcgc	ttccttttag	atgagctctt	gtaggatatg	ttgtcttctg	ctgcttgagt	6240
acagatccaa	ttgcagcata	tacttggcaa	gcagctcgta	caccgccttg	acagtttgcc	6300
ggcaacttgt	caatgecett	cttagctctc	accatgatag	cgtctgcctt	tcctacaagc	6360
ttcagagaca	agcccaacag	tctttgatca	cctaggctac	gggcgttgcc	ttgctgtatt	6420
tgttctgttt	cticttttct	aagccattgt	tgaggcagat	aacatcgacc	caacatcctc	6480
gagccatact	acagcataaa	aggatacgtt	ttctttaaca	gaaatttacc	cttttgttat	6540
cagcacatac	aaaaaaaaag	aaatttaaga	tgagtaggac	ttccattctc	tcaaaaattt	6600
tattcaatcc	ataaatgaat	tatttttgga	caaaaaagaa	agattatgcc	tgattttctc	6660
tattttttt	ttttttacaa	ctccaccaat	actttctagc	ccagcttggc	gtaatcatgg	6720
tcatagctgt	ttcctgtgtg	aaattgttat	ccgctcacaa	ttccacacaa	catacgagcc	6780
ggaagcataa	agtgtaaagc	ctggggtgcc	taatgagtga	gctaactcac	attaattgcg	6840
ttgcgctcac	tgeeegettt	ccagtcggga	aacctgtcgt	gccagctgca	ttaatgaatc	6900
ggccaacgcg	cggggagagg	cggtttgcgt	attgggccaa	agacaaaagg	gcgacattca	6960
accgattgag	ggagggaagg	taaatattga	cggaaattat	tcattaaagg	tgaattatca	7020
ccgtcaccga	cttgagccat	ttgggaatta	gagccagcaa	aatcaccagt	agcaccatta	7080
ccattagcaa	ggccggaaac	gtcaccaatg	aaaccatcga	tagcagcacc	gtaatcagta	7140
gcgacagaat	caagtttgcc	tttagcgtca	gactgtagcg	cgttttcatc	ggcattttcg	7200

WO 2004/063359 PCT/EP2004/000099 . 281/357

gtcatagccc	ccttattagc	gtttgccatc	ttttcataat	caaaatcacc	ggaaccagag	7260
ccaccaccgg	aaccgcctcc	ctcagagccg	ccaccctcag	aaccgccacc	ctcagagcca	7320
ccaccetcag	agccgccacc	agaaccacca	ccagagccgc	cgccagcatt	gacaggaggc	7380
ccgatctagt	aacatagatg	acaccgcgcg	cgataattta	tcctagtttg	cgcgctatat	7440
tttgttttct	atcgcgtatt	aaatgtataa	ttgcgggact	ctaatcataa	aaacccatct	7500
cataaataac	gtcatgcatt	acatgttaat	tattacatgc	ttaacgtaat	tcaacagaaa	7560
ttatatgata	atcatcgcaa	gaccggcaac	aggattcaat	cttaagaaac	tttattgcca	7620
aatgtttgaa	cgatcgggga	tcatccgggt	ctgtggcggg	aactccacga	aaatatccga	7680
acgcagcaag	atatcgcggt	gcatctcggt	cttgcctggg	cagtcgccgc	cgacgccgtt	7740
gatgtggacg	ccgggcccga	tcatattgtc	gctcaggatc	gtggcgttgt	gcttgtcggc	7800
cgttgctgtc	gtaatgatat	cggcaccttc	gaccgcctgt	tccgcagaga	tcccgtgggc	7860
gaagaactcc	agcatgagat	ccccgcgctg	gaggatcatc	cagccggcgt	cccggaaaac	7920
gattccgaag	cccaaccttt	catagaaggc	ggcggtggaa	tcgaaatctc	gtgatggcag	7980
gttgggcgtc	gcttggtcgg	tcatttcgaa	ccccagagtc	ccgctcagaa	gaactcgtca	8040
agaaggcgat	agaaggcgat	gcgctgcgaa	tcgggagcgg	cgataccgta	aagcacgagg	8100
aagcggtcag	cccattcgcc	gccaagctct	tcagcaatat	cacgggtagc	caacgctatg	8160
tcctgatagc	ggtccgccac	acccagccgg	ccacagtcga	tgaatccaga	aaagcggcca	8220
ttttccacca	tgatattcgg	caagcaggca	tcgccatggg	tcacgacgag	atcatcgccg	8280
tegggeatge	gcgccttgag	cctggcgaac	agttcggctg	gegegageed	ctgatgctct	8340
tcgtccagat	catcctgatc	gacaagaccg	gcttccatcc	gagtacgtgo	tegetegatg	8400
cgatgtttcg	cttggtggtc	gaatgggcag	gtagccggat	caagcgtatg	cageegeege	8460

WO 2004/063359 PCT/EP2004/000099 282/357

attgcatcag	ccatgatgga	tactttctcg	gcaggagcaa	ggtgagatga	caggagatcc	8520
tgccccggca	cttcgcccaa	tagcagccag	tcccttcccg	cttcagtgac	aacgtcgagc	8580
acagctgcgc	aaggaacgcc	cgtcgtggcc	agccacgata	gccgcgctgc	ctcgtcctgc	8640
agttcattca	gggcaccgga	caggtcggtc	ttgacaaaaa	gaaccgggcg	cccctgcgct	8700
gacagccgga	acacggcggc	atcagagcag	ccgattgtct	gttgtgccca	gtcatagccg	8760
aatagcctct	ccacccaagc	ggccggagaa	cctgcgtgca	atccatcttg	ttcaatcatg	8820
cgaaacgatc	cagatccggt	gcagattatt	tggattgaga	gtgaatatga	gactctaatt	8880
ggataccgag	gggaatttat	ggaacgtcag	tggagcattt	ttgacaagaa	atatttgcta	8940
gctgatagtg	accttaggcg	acttttgaac	gcgcaataat	ggtttctgac	gtatgtgctt	9000
agctcattaa	actccagaaa	cccgcggctg	agtggctcct	tcaacgttgc	ggttctgtca	9060
gttccaaacg	taaaacggct	tgtcccgcgt	catcggcggg	ggtcataacg	tgactccctt	9120
aattctccgc	tcatgatcag	attgtcgttt	cccgccttca	gtttaaacta	tcagtgtttg	9180
acaggatata	ttggcgggta	aacctaagag	aaaagagcgt	ttattagaat	aatcggatat	9240
ttaaaagggc	gtgaaaaggt	ttatccgttc	gtccatttgt	atgtgcatgc	caaccacagg	9300
gttccccaga	tetggegeeg	gccagcgaga	cgagcaagat	tggccgccgc	ccgaaacgat	9360
ccgacagcgc	gcccagcaca	ggtgcgcagg	caaattgcac	caacgcatac	agcgccagca	9420
gaatgccata	gtgggcggtg	acgtcgttcg	agtgaaccag	atcgcgcagg	aggcccggca	9480
gcaccggcat	aatcaggccg	atgccgacag	cgtcgagcgc	gacagtgctc	agaattacga	9540
tcaggggtat	gttgggtttc	acgtctggcc	tccggaccag	ceteegetgg	tccgattgaa	9600
cgcgcggatt	ctttatcact	gataagttgg	tggacatatt	: atgtttatca	gtgataaagt	9660
gtcaagcato	g acaaagttgc	agccgaatac	: agtgatccgt	gccgccctgg	g acctgttgaa	9720

WO 2004/063359 PCT/EP2004/000099 283/357

			•			
cgaggtcggc	gtagacggtc	tgacgacacg	caaactggcg	gaacggttgg	gggttcagca	9780
gccggcgctt	tactggcact	tcaggaacaa	gcgggcgctg	ctcgacgcac	tggccgaagc	9840
catgctggcg	gagaatcata	cgcattcggt	gccgagagcc	gacgacgact	ggcgctcatt	9900
tctgatcggg	aatgcccgca	gcttcaggca	ggcgctgctc	gcctaccgcg	atggcgcgcg	9960
catccatgcc	ggcacgcgac	cgggcgcacc	gcagatggaa	acggccgacg	cgcagcttcg	10020
cttcctctgc	gaggcgggtt	tttcggccgg	ggacgccgtc	aatgcgctga	tgacaatcag	10080
ctacttcact	gttggggccg	tgcttgagga	gcaggccggc	gacagcgatg	ccggcgagcg	10140
cggcggcacc	gttgaacagg	ctccgctctc	gccgctgttg	cgggccgcga	tagacgcctt	10200
cgacgaagcc	ggtccggacg	cagcgttcga	gcagggactc	gcggtgattg	tcgatggatt	10260
ggcgaaaagg	aggctcgttg	tcaggaacgt	tgaaggaccg	agaaagggtg	acgattgatc	10320
aggaccgctg	ccggagcgca	acccactcac	tacagcagag	ccatgtagac	aacatcccct	10380
cccctttcc	accgcgtcag	acgcccgtag	cagcccgcta	cgggcttttt	catgccctgc	10440
cctagcgtcc	aagcctcacg	geegegeteg	gcctctctgg	cggccttctg	gcgctcttcc	10500
gcttcctcgc	tcactgactc	gctgcgctcg	gtcgttcggc	tgcggcgagc	ggtatcagct	10560
cactcaaagg	cggtaatacg	gttatccaca	gaatcagggg	ataacgcagg	aaagaacatg	10620
tgagcaaaag	gccagcaaaa	ggccaggaac	cgtaaaaagg	ccgcgttgct	ggcgtttttc	10680
cataggctcc	gcccccctga	cgagcatcac	aaaaatcgac	gctcaagtca	gaggtggcga	10740
aacccgacag	gactataaag	ataccaggcg	tttccccctg	gaageteeet	cgtgcgctct	10800
cctgttccga	ccctgccgct	taccggatac	ctgtccgcct	ttetecette	gggaagcgtg	10860
gcgcttttcc	gctgcataac	cctgcttcgg	ggtcattata	gcgattttt	cggtatatcc	10920
atcctttttc	gcacgatata	caggattttg	ccaaagggtt	cgtgtagact	ttccttggtg	10980

WO 2004/063359 PCT/EP2004/000099 284/357

tatccaacgg	cgtcagccgg	gcaggatagg	tgaagtaggc	ccacccgcga	gcgggtgttc	11040
cttcttcact	gtcccttatt	cgcacctggc	ggtgctcaac	gggaatcctg	ctctgcgagg	11100
ctggccggct	accgccggcg	taacagatga	gggcaagcgg	atggctgatg	aaaccaagcc	11160
aaccaggaag	ggcagcccac	ctatcaaggt	gtactgcctt	ccagacgaac	gaagagcgat	11220
tgaggaaaag	gcggcggcgg	ccggcatgag	cctgtcggcc	tacctgctgg	ccgtcggcca	11280
gggctacaaa	atcacgggcg	tcgtggacta	tgagcacgtc	cgcgagctgg	cccgcatcaa	11340
tggcgacctg	ggccgcctgg	gcggcctgct	gaaactctgg	ctcaccgacg	acccgcgcac	11400
ggcgcggttc	ggtgatgcca	cgatcctcgc	cctgctggcg	aagatcgaag	agaagcagga	11460
cgagcttggc	aaggtcatga	tgggcgtggt	ccgcccgagg	gcagagccat	gactttttta	11520
gccgctaaaa	cggccggggg	gtgcgcgtga	ttgccaagca	cgtccccatg	cgctccatca	11580
agaagagcga	cttcgcggag	ctggtgaagt	acatcaccga	cgagcaaggc	aagaccgagc	11640
gcctttgcga	cgctcaccgg	gctggttgcc	ctcgccgctg	ggctggcggc	cgtctatggc	11700
cctgcaaacg	cgccagaaac	gccgtcgaag	ccgtgtgcga	gacaccgcgg	ccgccggcgt	11760
tgtggatacc	tcgcggaaaa	cttggccctc	actgacagat	gaggggcgga	cgttgacact	11820
tgaggggccg	actcacccgg	cgcggcgttg	acagatgagg	ggcaggctcg	atttcggccg	11880
gcgacgtgga	gctggccagc	ctcgcaaatc	ggcgaaaacg	cctgatttta	cgcgagtttc	11940
ccacagatga	tgtggacaag	cctggggata	agtgccctgc	ggtattgaca	cttgaggggc	12000
gcgactactg	acagatgagg	ggcgcgatcc	ttgacacttg	aggggcagag	tgctgacaga	12060
tgaggggcgc	acctattgac	atttgagggg	ctgtccacag	gcagaaaatc	cagcatttgc	12120
aagggtttcc	gcccgtttt	cggccaccgc	taacctgtct	tttaacctgc	ttttaaacca	12180
atatttataa	accttgtttt	taaccagggc	tgcgccctgt	gcgcgtgacc	gegeaegeeg	12240

WO 2004/063359 PCT/EP2004/000099 285/357

aaggggggtg	ccccccttc	tcgaaccete	ccggcccgct	aacgcgggcc	tcccatcccc	12300
ccaggggctg	cgcccctcgg	ccgcgaacgg	cctcacccca	aaaatggcag	cgctggcagt	12360
ccttgccatt	gccgggatcg	gggcagtaac	gggatgggcg	atcagcccga	gcgcgacgcc	12420
cggaagcatt	gacgtgccgc	aggtgctggc	atcgacattc	agcgaccagg	tgccgggcag	12480
tgagggcggc	ggcctgggtg	geggeetgee	cttcacttcg	gccgtcgggg	cattcacgga	12540
cttcatggcg	gggccggcaa	tttttacctt	gggcattctt	ggcatagtgg	tcgcgggtgc	12600
cgtgctcgtg	ttcgggggtg	cgataaaccc	agcgaaccat	ttgaggtgat	aggtaagatt	12660
ataccgaggt	atgaaaacga	gaattggacc	tttacagaat	tactctatga	agcgccatat	12720
ttaaaaagct	accaagacga	agaggatgaa	gaggatgagg	aggcagattg	ccttgaatat	12780
attgacaata	. ctgataagat	aatatatctt	ttatatagaa	gatatcgccg	tatgtaagga	12840
tttcaggggg	caaggcatag	gcagcgcgct	tatcaatata	tctatagaat	gggcaaagca	12900
taaaaacttg	r catggactaa	tgcttgaaac	ccaggacaat	aaccttatag	cttgtaaatt	12960
ctatcataat	: tgggtaatga	ctccaactta	ttgatagtgt	tttatgttca	gataatgccc	13020
gatgactttg	tcatgcagct	ccaccgattt	tgagaacgac	agcgacttcc	gtcccagccg	13080
tgccaggtgc	: tgcctcagat	tcaggttatg	ccgctcaatt	cgctgcgtat	atcgcttgct	13140
gattacgtgo	agctttccct	tcaggcggga	ttcatacagc	ggccagccat	ccgtcatcca	13200
tatcaccacg	g tcaaagggtg	acagcaggct	cataagacgc	cccagcgtcg	ccatagtgcg	13260
ttcaccgaat	: acgtgcgcaa	caaccgtctt	ccggagactg	tcatacgcgt	aaaacagcca	13320
gcgctggcgo	gatttagccc	cgacatagcc	ccactgttcg	tccatttccg	cgcagacgat	13380
gacgtcacts	g cccggctgta	tgcgcgaggt	taccgactgo	ggcctgagtt	tttaagtga	13440
cgtaaaatc	g tgttgaggco	aacgcccata	. atgcgggctg	ttgcccggca	tccaacgcca	13500

ttcatggcca	tatcaatgat	tttctggtgc	gtaccgggtt	gagaagcggt	gtaagtgaac	13560
tgcagttgcc	atgttttacg	gcagtgagag	cagagatagc	gctgatgtcc	ggcggtgctt	13620
ttgccgttac	gcaccacccc	gtcagtagct	gaacaggagg	gacagctgat	agacacagaa	13680
gccactggag	cacctcaaaa	acaccatcat	acactaaatc	agtaagttgg	cagcatcacc	13740
cataattgtg	gtttcaaaat	eggeteegte	gatactatgt	tatacgccaa	ctttgaaaac	13800
aactttgaaa	aagctgtttt	ctggtattta	aggttttaga	atgcaaggaa	cagtgaattg	13860
gagttcgtct	tgttataatt	agcttcttgg	ggtatcttta	aatactgtag	aaaagaggaa	13920
ggaaataata	aatggctaaa	atgagaatat	caccggaatt	gaaaaaactg	atcgaaaaat	13980
accgctgcgt	aaaagatacg	gaaggaatgt	ctcctgctaa	ggtatataag	ctggtgggag	14040
aaaatgaaaa	cctatattta	aaaatgacgg	acagccggta	taaagggacc	acctatgatg	14100
tggaacggga	aaaggacatg	atgctatggc	tggaaggaaa	gctgcctgtt	ccaaaggtcc	14160
tgcactttga	acggcatgat	ggctggagca	atctgctcat	gagtgaggcc	gatggcgtcc	14220
tttgctcgga	agagtatgaa	gatgaacaaa	gccctgaaaa	gattatcgag	ctgtatgcgg	14280
agtgcatcag	gctctttcac	tccatcgaca	tatcggattg	tccctatacg	aatagcttag	14340
acagccgctt	agccgaattg	gattacttac	tgaataacga	tctggccgat	gtggattgcg	14400
aaaactggga	agaagacact	ccatttaaag	atccgcgcga	gctgtatgat	tttttaaaga	14460
cggaaaagco	cgaagaggaa	. cttgtcttt	cccacggcga	cctgggagac	agcaacatct	14520
ttgtgaaaga	tggcaaagta	agtggcttta	ttgatcttgg	gagaagcggc	agggcggaca	14580
agtggtatga	cattgcctto	: tgcgtccggt	cgatcaggga	ggatateggg	gaagaacagt	14640
atgtcgagct	: attttttgac	: ttactgggga	ı tcaagcctga	ı ttgggagaaa	ataaaatatt	14700
atattttact	ggatgaattg	ttttagtaco	: tagatgtggc	gcaacgatgo	: cggcgacaag	14760

WO 2004/063359 PCT/EP2004/000099 287/357

caggagcgca	ccgacttctt	ccgcatcaag	tgttttggct	ctcaggccga	ggeceaegge	14820
aagtatttgg	gcaaggggtc	gctggtattc	gtgcagggca	agattcggaa	taccaagtac	14880
gagaaggacg	gccagacggt	ctacgggacc	gacttcattg	ccgataaggt	ggattatctg	14940
gacaccaagg	caccaggcgg	gtcaaatcag	gaataagggc	acattgcccc	ggcgtgagtc	15000
ggggcaatcc	cgcaaggagg	gtgaatgaat	cggacgtttg	accggaaggc	atacaggcaa	15060
gaactgatcg	acgcggggtt	ttccgccgag	gatgccgaaa	ccatcgcaag	ccgcaccgtc	15120
atgcgtgcgc	cccgcgaaac	cttccagtcc	gtcggctcga	tggtccagca	agctacggcc	15180
aagatcgagc	gcgacagcgt	gcaactggct	cccctgccc	tgcccgcgcc	ateggeegee	15240
gtggagcgtt	cgcgtcgtct	cgaacaggag	geggeaggtt	tggcgaagtc	gatgaccatc	15300
gacacgcgag	gaactatgac	gaccaagaag	cgaaaaaccg	ccggcgagga	cctggcaaaa	15360
caggtcagcg	aggccaagca	ggccgcgttg	ctgaaacaca	cgaagcagca	gatcaaggaa	15420
atgcagcttt	ccttgttcga	tattgcgccg	tggccggaca	cgatgcgagc	gatgccaaac	15480
gacacggccc	getetgeeet	gttcaccacg	cgcaacaaga	aaatcccgcg	cgaggcgctg	15540
caaaacaagg	tcattttcca	cgtcaacaag	gacgtgaaga	tcacctacac	cggcgtcgag	15600
ctgcgggccg	acgatgacga	actggtgtgg	cagcaggtgt	tggagtacgc	gaagegeace	15660
cctatcggcg	agccgatcac	cttcacgttc	tacgagcttt	gccaggacct	gggctggtcg	15720
atcaatggco	ggtattacac	gaaggccgag	gaatgcctgt	cgcgcctaca	ggcgacggcg	15780
atgggcttca	cgtccgaccg	· cgttgggcac	ctggaatcgg	tgtegetget	gcaccgcttc	15840
cgcgtcctgg	accgtggcaa	gaaaacgtco	cgttgccagg	tcctgatcga	cgaggaaatc	15900
gtcgtgctgt	: ttgctggcga	ccactacacg	aaattcatat	: gggagaagta	ccgcaagctg	15960
tcgccgacgg	g cccgacggat	gttcgactat	ttcagctcgc	accgggagcc	gtacccgctc	16020

aagctggaaa ccttccgcct catgtgcgga tcggattcca cccgcgtgaa gaagtggcgc 1608	В0
gagcaggtcg gcgaagcctg cgaagagttg cgaggcagcg gcctggtgga acacgcctgg 1614	40
gtcaatgatg acctggtgca ttgcaaacgc tagggccttg tggggtcagt tccggctggg 1620	00
ggttcagcag ccagcgcttt actggcattt caggaacaag cgggcactgc tcgacgcact 1620	60
tgettegete agtategete gggaegeaeg gegegeteta egaaetgeeg ataaacagag 1632	20
gattaaaatt gacaattgtg attaaggctc agattcgacg gcttggagcg gccgacgtgc 163	80
aggatttccg cgagatccga ttgtcggccc tgaagaaagc tccagagatg ttcgggtccg 164	40
tttacgagca cgaggagaaa aagcccatgg aggcgttcgc tgaacggttg cgagatgccg 165	00
tggcattcgg cgcctacatc gacggcgaga tcattgggct gtcggtcttc aaacaggagg 165	60
acggccccaa ggacgctcac aaggcgcatc tgtccggcgt tttcgtggag cccgaacagc 166	20
gaggccgagg ggtcgccggt atgctgctgc gggcgttgcc ggcgggttta ttgctcgtga 166	80
tgategteeg acagatteea aegggaatet ggtggatgeg catetteate eteggegeae 167	40
ttaatatttc gctattctgg agcttgttgt ttatttcggt ctaccgcctg ccgggcgggg 168	100
tegeggegae ggtaggeget gtgeageege tgatggtegt gtteatetet geegetetge 168	360
taggtagccc gatacgattg atggcggtcc tgggggctat ttgcggaact gcgggcgtgg 169	€20
cgctgttggt gttgacacca aacgcagcgc tagatcctgt cggcgtcgca gcgggcctgg 169	980
cggggggggt ttccatggcg ttcggaaccg tgctgacccg caagtggcaa cctcccgtgc 170)40
ctctgctcac ctttaccgcc tggcaactgg cggccggagg acttctgctc gttccagtag 173	100
ctttagtgtt tgatccgcca atcccgatgc ctacaggaac caatgttctc ggcctggcgt 173	160
ggctcggcct gatcggagcg ggtttaacct acttcctttg gttccggggg atctcgcgac 173	220
tcgaacctac agttgtttcc ttactgggct ttctcagccc cagatctggg gtcgatcagc 172	280

WO 2004/063359 PCT/EP2004/000099 289/357

cggggatgca	tcaggccgac	agtcggaact	tegggteece	gacctgtacc	attcggtgag	17340
caatggatag	gggagttgat	atcgtcaacg	ttcacttcta	aagaaatagc	gccactcagc	17400
ttcctcagcg	gctttatcca	gcgatttcct	attatgtcgg	catagttctc	aagatcgaca	17460
gcctgtcacg	gttaagcgag	aaatgaataa	gaaggctgat	aattcggatc	tctgcgaggg	17520
agatgatatt	tgatcacagg	cagcaacgct	ctgtcatcgt	tacaatcaac	atgctaccct	17580
ccgcgagatc	atccgtgttt	caaacccggc	agcttagttg	ccgttcttcc	gaatagcatc	17640
ggtaacatga	gcaaagtctg	ccgccttaca	acggctctcc	cgctgacgcc	gtcccggact	17700
gatgggctgc	ctgtatcgag	tggtgatttt	gtgccgagct	gccggtcggg	gagctgttgg	17760
ctggctggtg	gcaggatata	ttgtggtgta	aacaaattga	cgcttagaca	acttaataac	17820
acattgcgga	cgtttttaat	gtactggggt	ggtttttctt	ttcaccagtg	agacgggcaa	17880
cagctgattg	cccttcaccg	cctggccctg	agagagttgc	agcaagcggt	ccacgctggt	17940
ttgccccagc	aggcgaaaat	cctgtțtgat	ggtggttccg	aaatcggcaa	aatcccttat	18000
aaatcaaaag	aatagcccga	gatagggttg	agtgttgttc	cagtttggaa	caagagtcca	18060
ctattaaaga	acgtggactc	caacgtcaaa	gggcgaaaaa	ccgtctatca	gggcgatggc	18120
ccactacgtg	aaccatcacc	caaatcaagt	tttttggggt	cgaggtgccg	taaagcacta	18180
aatcggaacc	ctaaagggag	cccccgattt	agagettgae	ggggaaagcc	ggcgaacgtg	18240
gcgagaaagg	aagggaagaa	agcgaaagga	gegggegeea	ttcaggctgc	gcaactgttg	18300
ggaagggcga	tcggtgcggg	cctcttcgct	attacgccag	ctggcgaaag	ggggatgtgc	18360
tgcaaggcga	ttaagttggg	taacgccagg	gttttcccag	tcacgacgtt	gtaaaacgac	18420
ggccagtgaa	ttcgagctcg	gtacccggg				18449

290/357 <210> 50 <211> 18617 <212> DNA <213> Artificial <220> <223> Plasmid <220> <221> misc_feature <222> (10264)..(10264) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10472)..(10472) <223> n is a, c, g, or t <220> <221> misc_feature <222> (10563)..(10563) <223> n is a, c, g, or t <400> 50 cegggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60 120 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 180 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 240 300 cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 360 caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat 420 tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480

ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg

540

WO 2004/063359 PCT/EP2004/000099 291/357

tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcgcac	gccgaagggg	ggtgcccccc	600
cttctcgaac	cctcccggcc	cgctaacgcg	ggcctcccat	cccccaggg	gctgcgcccc	660
teggeegega	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	cattgccggg	720
atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	cattgacgtg	780
ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	cggcggcctg	840
ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	ggcggggccg	900
gcaatttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	cgtgttcggg	960
ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	aggtatgaaa	1020
acgagaattg	gacctttaca	gaattactct	atgaagcgcc	atatttaaaa	agctaccaag	1080
acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	aatactgata	1140
agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	ggggcaaggc	1200
ataggcagcg	cgcttatcaa	tatatctata	gaatgggcaa	agcataaaaa	cttgcatgga	1260
ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgccgctc	aattcgctgc	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cagcggccag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg	tcttccggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat	agccccactg	ttcgtccatt	tccgcgcaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg	aggttaccga	ctgcggcctg	agtttttaa	gtgacgtaaa	atcgtgttga	1800

WO 2004/063359 PCT/EP2004/000099 292/357

ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
						1020
tgattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg	agagcagaga	tagcgctgat	gteeggeggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc	2040
aaaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc	cgtcgatact	atgttatacg	ccaactttga	aaacaacttt	gaaaaagctg	2160
ttttctggta	tttaaggttt	tagaatgcaa	ggaacagtga	attggagttc	gtcttgttat	2220
aattagcttc	ttggggtatc	tttaaatact	gtagaaaaga	ggaaggaaat	aataaatggc	2280
taaaatgaga	atatcaccgg	aattgaaaaa	actgatcgaa	aaataccgct	gcgtaaaaga	2340
tacggaagga	atgteteetg	ctaaggtata	taagctggtg	ggagaaaatg	aaaacctata	2400
tttaaaaatg	acggacagcc	ggtataaagg	gaccacctat	gatgtggaac	gggaaaagga	2460
catgatgcta	tggctggaag	gaaagctgcc	tgttccaaag	gtcctgcact	ttgaacggca	2520
tgatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac	ttactgaata	acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt	aaagatccgc	gcgagctgta	tgatttttta	aagacggaaa	agcccgaaga	2820
ggaacttgtc	ttttcccacg	gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc	tttattgatc	ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc	cggtcgatca	gggaggatat	cggggaagaa	. cagtatgtcg	agctatttt	3000
tgacttactg	gggatcaagc	ctgattggga	gaaaataaaa	tattatatt	tactggatga	3060

WO 2004/063359 PCT/EP2004/000099 293/357

attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag cgcaccgact 3120 tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat ttgggcaagg 3180 ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag gacggccaga cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc aaggcaccag 3300 gegggtcaaa teaggaataa gggcacattg eeceggegtg agteggggca ateeegcaag 3360 gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg atcgacgcgg 3420 ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt gcgccccgcg 3480 aaaccttcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc gagcgcgaca 3540 gegtgeaact ggeteeect geeetgeeeg egeeategge egeegtggag egttegegte 3600 gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg cgaggaacta 3660 tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc agcgaggcca 3720 agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag ctttccttgt 3780 togatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg gcccgctctg 3840 3900 ccctgttcac cacgcgcaac aagaaaatcc cgcgcgaggc gctgcaaaac aaggtcattt tocacgtcaa caaggacgtg aagatcacct acaccggcgt cgagctgcgg gccgacgatg 3960 4020 acgaactggt gtggcagcag gtgttggagt acgcgaagcg cacccctatc ggcgagccga tcaccttcac gttctacgag ctttgccagg acctgggctg gtcgatcaat ggccggtatt 4080 acacgaagge egaggaatge etgtegegee tacaggegae ggegatggge tteaegteeg 4140 accgegttgg gcacctggaa teggtgtege tgctgcaccg etteegegte etggaccgtg 4200 4260 gcaagaaaac gtcccgttgc caggtcctga tcgacgagga aatcgtcgtg ctgtttgctg gcgaccacta cacgaaattc atatgggaga agtaccgcaa gctgtcgccg acggcccgac 4320

WO 2004/063359 PCT/EP2004/000099 . 294/357

,	ggatgttcga	ctatttcagc	tcgcaccggg	agccgtaccc	gctcaagctg	gaaaccttcc	4380
,	gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
	cctgcgaaga	gttgcgaggc	agcggcctgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
	tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	gcagccagcg	4560
	ctttactggc	atttcaggaa	caagcgggca	ctgctcgacg	cacttgcttc	gctcagtatc	4620
	gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
	tgtgattaag	gctcagattc	, gacggcttgg	agcggccgac	gtgcaggatt	tccgcgagat	4740
	ccgattgtcg	gccctgaaga	aagctccaga	gatgttcggg	tccgtttacg	agcacgagga	4800
	gaaaaagccc	atggaggcgt	tcgctgaacg	gttgcgagat	gccgtggcat	teggegeeta	4860
	catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
	tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagegaggee	gaggggtcgc	4980
	cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	tccgacagat	5040
	tccaacggga	atctggtgga	tgcgcatctt	catcctcggc	gcacttaata	tttcgctatt	5100
	ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	cgacggtagg	5160
	cgctgtgcag	ccgctgatgg	tcgtgttcat	ctctgccgct	ctgctaggta	gcccgatacg	5220
	attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	tggtgttgac	5280
	accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	cggtttccat	5340
	ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	tcacctttac	5400
	cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	tgtttgatcc	5460
	gccaatcccg	atgcctacag	gaaccaatgt	tctcggcctg	gcgtggctcg	gcctgatcgg	5520
	agcgggttta	acctacttcc	tttggttccg	ggggatctcg	cgactcgaac	ctacagttgt	5580

WO 2004/063359 PCT/EP2004/000099 295/357

tteettactg ggetttetea geeceagate tggggtegat eageegggga tgeateagge 5640 cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5700 tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5760 tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5820 cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca 5880 caggcagcaa cgctctgtca tcgttacaat caacatgcta ccctccgcga gatcatccgt 5940 gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000 6060 totgoogcot tacaacggot otcoogctga cgccgtcccg gactgatggg ctgcctgtat 6120 cgagtggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga 6180 tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240 6300 accgcctggc cctgagagag ttgcagcaag cggtccacgc tggtttgccc cagcaggcga 6360 aaatcctgtt tgatggtggt tccgaaatcg gcaaaatccc ttataaatca aaagaatagc 6420 ccgagatagg gttgagtgtt gttccagttt ggaacaagag tccactatta aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta cgtgaaccat 6480 6540 cacccaaatc aagttttttg gggtcgaggt gccgtaaagc actaaatcgg aaccctaaag 6600 ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga aaggaaggga agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg gcgatcggtg 6660 6720 cgggcctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag gcgattaagt 6780 tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag tgaattcgag ctcggtaccc ggggatcttt cgacactgaa atacgtcgag cctgctccgc ttggaagcgg 6840

WO 2004/063359 PCT/EP2004/000099 296/357

cgaggagcct	cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900
gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac	gctcttctga	tccagtcgat	catccgctga	aggcgctttc	gaatctggtt	7020
aagatccacg	tcttcgggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140
aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag	aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320
tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
tcgaaataca	tccggatgtc	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560
tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaacatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ggtggagagc	ttataccgag	ctcccaaatc	7920
tgtccagatc	atggttgacc	ggtgcctgga	tcttcctata	gaatcatcct	tattcgttga	7980
cctagctgat	tctggagtga	cccagagggt	catgacttga	gcctaaaatc	cgccgcctcc	8040
accatttgta	gaaaaatgtg	acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100

WO 2004/063359 PCT/EP2004/000099 297/357

tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataagc	cactggccag	8160
acagctctgg	cggctctgag	gtgcagtgga	tgattattaa	teegggaeeg	geegeeete	8220
cgccccgaag	tggaaaggct	ggtgtgcccc	tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agecaggggt	8340
gtatagccgt	cggcgaaata	gcatgccatt	aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	agattcacga	gatagtacct	tctccgaagt	aggtagagcg	agtacccggc	8460
gcgtaagctc	cctaattggc	ccatccggca	tctgtagggc	gtccaaatat	egtgeetete	8520
ctgctttgcc	cggtgtatga	aaccggaaag	geegeteagg	agctggccag	cggcgcagac	8580
cgggaacaca	agctggcagt	cgacccatcc	ggtgctctgc	actcgacctg	ctgaggtccc	8640
tcagtccctg	gtaggcagct	ttgccccgtc	tgtccgcccg	gtgtgtcggc	ggggttgaca	8700
aggtcgttgc	gtcagtccaa	catttgttgc	catattttcc	tgctctcccc	accagctgct	8760 ·
cttttcttt	ctctttcttt	teccatette	agtatattca	tetteceate	caagaacctt	8820
tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgo	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggaggg	gtggatatgt	cctgcgggta	. aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gegetecega	ttccggaagt	9180
gcttgacatt	: ggggaattca	ı gegagageet	gacctattgo	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	r caagacctgo	ctgaaaccga	actgcccgct	gttctgcago	: cggtcgcgga	9300
ggccatggat	gcgatcgctg	g cggccgatct	tagccagacg	agcgggttcg	gcccattcgg	9360

WO 2004/063359 PCT/EP2004/000099 298/357

accgcaagga	atcggtcaat	acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	tegegeagge	9480
tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	tegtgeaege	9540
ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	cggagcttgc	9720
aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga	gccgggactg	tegggegtae	acaaatcgcc	cgcagaagcg	eggeegtetg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960
gagggcaaag	gaatagagta	gatgccgacc	gcgggatcga	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agatcgttgg	tgtcgatgtc	agctccggag	10080
ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttcgggttta	10200
cctcttccag	atacagctca	tetgcaatge	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagago	tattttcatt	ttcgggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380
tccacgcgac	tatatattt <u>c</u>	, tctctaattg	g tactttgaca	ı tgeteetett	ctttactctg	10440
					attgcatgtt	10500
					a acctcgaaat	10560
canttectac	: taagatggta	a tacaatagta	a accatgcat	g gttgcctagt	gaatgctccg	10620

WO 2004/063359 PCT/EP2004/000099 299/357

taacacccaa tacgccggcc gaaacttttt tacaactctc ctatgagtcg tttacccaga 10680 atgcacaggt acacttgttt agaggtaatc cttctttcta gctagaagtc ctcgtgtact 10740 gtgtaagcgc ccactccaca tctccactcg acctgcaggc atgcaagctt gagattaaaa 10800 tagataagga aaagaaagtg aaaagaaatt cggaagcatg gcacattctt ctttttataa 10860 atacatgcct gactttcttt ttccatcgat atgatatatg catatgatag atatacaagc 10920 aatcttcttc aaggagtttg aaattttgtc ctccaggagc aaaaaaaagt tttttttat 10980 acatgtttgt acacaagaat agttaccaat ttgctttggt cttacgtgct gcaagtttat 11040 atogtttca atttctttgt ctttacattt tctttgtcct ttatctttcc tcatttagtc 11100 tttgggagaa ttaggaaaag ggagcggaaa ggtaagaaat gcttgcgtat tttactaatt 11160 cggcaaacat ccaatttggc aaacagcagc ctgtgcaacg ctctcgagat gacagtatct 11220 ttgattacac tctaaatctc gatgacccga ccaaaaagag cgaacaaaga aataatcttg 11280 tgcattcgaa tatgatggaa gattttttcc cccttattct aaatgttgac atagcgtgta 11340 tgttatataa acaaaaagaa attgtacaaa ctttcttttc ttctctttt attttatctc 11400 tatgctgtcg aagctgcagt caatcagcgt caaggcccgc cgcgttgaac tagcccgcga 11460 catcacgcgg cccaaagtct gcctgcatgc tcagcggtgc tcgttagttc ggctgcgagt 11520 ggcagcacca cagacagagg aggcgctggg aaccgtgcag gctgccggcg cgggcgatga 11580 gcacagegee gatgtageae tecageaget tgaceggget ategeagage gtegtgeeeg 11640 gegeaaaegg gageagetgt cataceagge tgeegecatt geageateaa ttggegtgte 11700 aggeattgee atettegeea cetacetgag atttgecatg cacatgaceg tgggeggege 11760 agtgccatgg ggtgaagtgg ctggcactct cctcttggtg gttggtggcg cgctcggcat 11820 ggagatgtat gcccgctatg cacacaaagc catctggcat gagtcgcctc tgggctggct 11880

WO 2004/063359 PCT/EP2004/000099 300/357

gctgcacaag	agccaccaca	cacctcgcac	tggacccttt	gaagccaacg	acttgtttgc	11940
aatcatcaat	ggactgcccg	ccatgctcct	gtgtaccttt	ggcttctggc	tgcccaacgt	12000
cctgggggcg	gcctgctttg	gagcggggct	gggcatcacg	ctatacggca	tggcatatat	12060
gtttgtacac	gatggcctgg	tgcacaggcg	ctttcccacc	gggcccatcg	ctggcctgcc	12120
ctacatgaag	cgcctgacag	tggcccacca	gctacaccac	agcggcaagt	acggtggcgc	12180
gccctggggt	atgttcttgg	gtccacagga	getgeageae	attccaggtg	cggcggagga	12240
ggtggagcga	ctggtcctgg	aactggactg	gtccaagcgg	tagaagcttg	agattaaaat	12300
agataaggaa	aagaaagtga	aaagaaattc	ggaagcatgg	cacattcttc	tttttataaa	12360
tacatgcctg	actttcttt	tccatcgata	tgatatatgc	atatgataga	tatacaagca	12420
atcttcttca	aggagtttga	aattttgtcc	tccaggagca	aaaaaaagtt	ttttttata	12480
catgtttgta	cacaagaata	gttaccaatt	tgctttggtc	ttacgtgctg	caagtttata	12540
tcgttttcaa	tttctttgtc	tttacatttt	ctttgtcctt	tatctttcct	catttagtct	12600
ttgggagaat	taggaaaagg	gagcggaaag	gtaagaaatg	cttgcgtatt	ttactaattc	12660
ggcaaacatc	caatttggca	aacagcagcc	tgtgcaacgc	tctcgagatg	acagtatett	12720
tgattacact	ctaaatctcg	atgacccgac	caaaaagagc	gaacaaagaa	ataatcttgt	12780
gcattcgaat	atgatggaag	attttttccc	ccttattcta	. aatgttgaca	tagcgtgtat	12840
gttatataaa	. caaaaagaaa	ttgtacaaac	tttcttttct	tetetttta	tttatctct	12900
atgatccagt	tagaacaacc	actcagtcat	caagcaaaac	: tgactccagt	actgagaagt	12960
aaatctcagt	ttaaggggct	tttcattgct	attgtcattg	, ttagcgcatc	ggtcattagc	13020
ctgagtttat	: tactttccct	: tgacatctca	aagctaaaat	: tttggatgtt	attgcctgtt	13080
atactatggo	aaacattttt	: atatacggga	ttatttatta	a catctcatga	a tgccatgcat	13140

WO 2004/063359 PCT/EP2004/000099 301/357

ggcgtagtat	ttccccaaaa	caccaagatt	aatcatttga	ttggaacatt	gaccctatcc	13200
ctttatggtc	ttttaccata	tcaaaaacta	ttgaaaaaac	attggttaca	ccaccacaat	13260
ccagcaagct	caatagaccc	ggattttcac	aatggtaaac	accaaagttt	ctttgcttgg	13320
tattttcatt	ttatgaaagg	ttactggagt	tgggggcaaa	taattgcgtt	gactattatt	13380
tataactttg	ctaaatacat	actccatatc	ccaagtgata	atctaactta	cttttgggtg	13440
ctaccctcgc	tttaagttc	attacaatta	ttctattttg	gtacttttt	accccatagt	13500
gaaccaatag	ggggttatgt	tcagcctcat	tgtgcccaaa	caattagccg	tcctatttgg	13560
tggtcattta	tcacgtgcta	tcattttggc	taccacgagg	aacatcacga	atatcctcat	13620
atttcttggt	ggcagttacc	agaaatttac	aaagcaaaat	agaagcttgg	cgtaatcatg	13680
gtcatagctg	tttcctgtgt	gaaattgtta	tccgctcaca	attccacaca	acatacgagc	13740
cggaagcata	aagtgtaaag	cctggggtgc	ctaatgagtg	agctaactca	cattaattgc	13800
gttgcgctca	ctgcccgctt	tccagtcggg	aaacctgtcg	tgccagctgc	attaatgaat	13860
cggccaacgc	gcggggagag	gcggtttgcg	tattgggcca	aagacaaaag	ggcgacattc	13920
aaccgattga	gggagggaag	gtaaatattg	acggaaatta	ttcattaaag	gtgaattatc	13980
accgtcaccg	acttgagcca	tttgggaatt	agagccagca	aaatcaccag	tagcaccatt	14040
accattagca	aggccggaaa	cgtcaccaat	gaaaccatcg	atagcagcac	cgtaatcagt	14100
agcgacagaa	tcaagtttgc	ctttagcgtc	agactgtagc	gcgttttcat	cggcattttc	14160
ggtcatagcc	cccttattag	cgtttgccat	cttttcataa	tcaaaatcac	cggaaccaga	14220
gccaccaccg	gaaccgcctc	cctcagagcc	gccaccctca	gaaccgccac	cctcagagcc	14280
accaccctca	gagccgccac	cagaaccacc	accagagccg	ccgccagcat	tgacaggagg	14340
cccgatctag	taacatagat	gacacegege	gcgataattt	atcctagttt	gcgcgctata	14400

WO 2004/063359 PCT/EP2004/000099 302/357

ttttgttttc	tatcgcgtat	taaatgtata	attgcgggac	tctaatcata	aaaacccatc	14460
tcataaataa	cgtcatgcat	tacatgttaa	ttattacatg	cttaacgtaa	ttcaacagaa	14520
attatatgat	aatcatcgca	agaccggcaa	caggattcaa	tcttaagaaa	ctttattgcc	14580
aaatgtttga	acgatcgggg	atcatccggg	tetgtggegg	gaactccacg	aaaatatccg	14640
aacgcagcaa	gatatcgcgg	tgcatctcgg	tettgeetgg	gcagtcgccg	ccgacgccgt	14700
tgatgtggac	gccgggcccg	atcatattgt	cgctcaggat	cgtggcgttg	tgcttgtcgg	14760
ccgttgctgt	cgtaatgata	teggeacett	cgaccgcctg	ttccgcagag	atcccgtggg	14820
cgaagaactc	cagcatgaga	teccegeget	ggaggatcat	ccagccggcg	tcccggaaaa	14880
cgattccgaa	gcccaacctt	tcatagaagg	cggcggtgga	atcgaaatct	cgtgatggca	14940
ggttgggcgt	cgcttggtcg	gtcatttcga	accccagagt	cccgctcaga	agaactcgtc	15000
aagaaggcga	tagaaggcga	tgcgctgcga	atcgggagcg	gcgataccgt	aaagcacgag	15060
gaagcggtca	gcccattcgc	cgccaagctc	ttcagcaata	tcacgggtag	ccaacgctat	15120
gtcctgatag	cggtccgcca	cacccagccg	gccacagtcg	atgaatccag	aaaageggee	15180
attttccacc	atgatattcg	gcaagcaggc	atcgccatgg	gtcacgacga	gatcatcgcc	15240
gtcgggcatg	cgcgccttga	gcctggcgaa	cagttcggct	ggcgcgagcc	cctgatgctc	15300
ttcgtccaga	tcatcctgat	cgacaagacc	ggcttccatc	cgagtacgtg	ctcgctcgat	15360
gcgatgtttc	gcttggtggt	cgaatgggca	ggtagccgga	tcaagcgtat	gcagccgccg	15420
cattgcatca	gccatgatgg	atactttctc	ggcaggagca	aggtgagatg	acaggagatc	15480
ctgccccggc	acttcgccca	atagcagcca	gtecettece	gcttcagtga	caacgtcgag	15540
cacagetgeg	caaggaacgc	ccgtcgtggc	cagccacgat	ageegegetg	cctcgtcctg	15600
cagttcattc	agggcaccgg	acaggtcggt	cttgacaaaa	agaaccgggc	geecetgege	15660

WO 2004/063359 PCT/EP2004/000099 303/357

tgacagccgg	aacacggcgg	catcagagca	gccgattgtc	tgttgtgccc	agtcatagec	15720
gaatagcctc	tccacccaag	cggccggaga	acctgcgtgc	aatccatctt	gttcaatcat	15780
gcgaaacgat	ccagatccgg	tgcagattat	ttggattgag	agtgaatatg	agactctaat	15840
tggataccga	ggggaattta	tggaacgtca	gtggagcatt	tttgacaaga	aatatttgct	15900
agctgatagt	gaccttaggc	gacttttgaa	cgcgcaataa	tggtttctga	cgtatgtgct	15960
tagctcatta	aactccagaa	accegegget	gagtggctcc	ttcaacgttg	cggttctgtc	16020
agttccaaac	gtaaaacggc	ttgtcccgcg	tcatcggcgg	gggtcataac	gtgactccct	16080
taattctccg	ctcatgatca	gattgtcgtt	tecegeette	agtttaaact	atcagtgttt	16140
gacaggatat	attggcgggt	aaacctaaga	gaaaagagcg	tttattagaa	taatcggata	16200
tttaaaaggg	cgtgaaaagg	tttatccgtt	cgtccatttg	tatgtgcatg	ccaaccacag	16260
ggttccccag	atctggcgcc	ggccagcgag	acgagcaaga	ttggccgccg	cccgaaacga	16320
tccgacagcg	cgcccagcac	aggtgcgcag	gcaaattgca	ccaacgcata	cagegeeage	16380
agaatgccat	agtgggcggt	gacgtcgttc	gagtgaacca	gatcgcgcag	gaggcccggc	16440
agcaccggca	taatcaggcc	gatgccgaca	gcgtcgagcg	cgacagtgct	cagaattacg	16500
atcaggggta	tgttgggttt	cacgtctggc	ctccggacca	gcctccgctg	gtccgattga	16560
acgcgcggat	tetttateae	tgataagttg	gtggacatat	tatgtttatc	agtgataaag	16620
tgtcaagcat	gacaaagttg	cageegaata	cagtgatccg	tgccgccctg	gacctgttga	16680
acgaggtcgg	cgtagacggt	ctgacgacac	gcaaactggc	ggaacggttg	ggggttcagc	16740
agccggcgct	ttactggcac	ttcaggaaca	agcgggcgct	gctcgacgca	ctggccgaag	16800
ccatgctggc	ggagaatcat	acgcattcgg	tgccgagagc	cgacgacgac	tggcgctcat	16860
ttctgatcgg	gaatgcccgc	agcttcaggo	aggcgctgct	cgcctaccgc	gatggcgcgc	16920

WO 2004/063359 PCT/EP2004/000099 304/357

gcatccatgc	caacacacaa	ccaaacacac	cocagatoga	aacqqccqac	gcgcagcttc	16980
gcacccacgc	cggcacgcga	oogggogouo	og cag a c g g s			
gcttcctctg	cgaggcgggt	ttttcggccg	gggacgccgt	caatgcgctg	atgacaatca	17040
gctacttcac	tgttggggcc	gtgcttgagg	agcaggccgg	cgacagcgat	gccggcgagc	17100
geggeggeae	cgttgaacag	gctccgctct	cgccgctgtt	gcgggccgcg	atagacgcct	17160
tcgacgaagc	cggtccggac	gcagcgttcg	agcagggact	cgcggtgatt	gtcgatggat	17220
tggcgaaaag	gaggctcgtt	gtcaggaacg	ttgaaggacc	gagaaagggt	gacgattgat	17280
caggaccgct	gccggagcgc	aacccactca	ctacagcaga	gccatgtaga	caacatcccc	17340
teceeettte	caccgcgtca	gacgcccgta	gcagcccgct	acgggctttt	tcatgccctg	17400
ccctagcgtc	caageeteae	ggccgcgctc	ggcctctctg	gcggccttct	ggcgctcttc	17460
cgcttcctcg	ctcactgact	cgctgcgctc	ggtcgttcgg	ctgcggcgag	cggtatcagc	17520
tcactcaaag	gcggtaatac	ggttatccac	agaatcaggg	gataacgcag	gaaagaacat	1 7580
gtgagcaaaa	ggccagcaaa	aggccaggaa	ccgtaaaaag	geegegttge	tggcgttttt	17640
ccataggctc	cgcccccctg	acgagcatca	caaaaatcga	cgctcaagtc	agaggtggcg	17700
aaacccgaca	ggactataaa	gataccaggc	gtttcccct	ggaagctccc	tegtgegete	17760
tcctgttccg	accctgccgc	ttaccggata	cctgtccgcc	tttctccctt	cgggaagcgt	17820
ggcgcttttc	cgctgcataa	ccctgcttcg	gggtcattat	agcgattttt	teggtatate	17880
catcetttt	cgcacgatat	acaggatttt	gccaaagggt	tcgtgtagac	tttccttggt	17940
gtatccaacg	gcgtcagccg	ggcaggatag	gtgaagtagg	cccacccgcg	agcgggtgtt	18000
ccttcttcac	tgtcccttat	tegeacetgg	cggtgctcaa	cgggaatcct	gctctgcgag	18060
gctggccggc	taccgccggc	gtaacagatg	agggcaagcg	gatggctgat	gaaaccaagc	18120
caaccaggaa	gggcagccca	cctatcaagg	tgtactgcct	tccagacgaa	cgaagagcga	18180

ttgaggaaaa (ggcggcggcg	gccggcatga	gcctgtcggc	ctacctgctg	gccgtcggcc	18240
agggctacaa a	aatcacgggc	gtcgtggact	atgagcacgt	ccgcgagctg	gcccgcatca	18300
atggcgacct (gggccgcctg	ggcggcctgc	tgaaactctg	gctcaccgac	gacccgcgca	18360
cggcgcggtt	cggtgatgcc	acgatcctcg	ecctgctggc	gaagatcgaa	gagaagcagg	18420
acgagcttgg	caaggtcatg	atgggcgtgg	tccgcccgag	ggcagagcca	tgacttttt	18480
agccgctaaa a	acggccgggg	ggtgcgcgtg	attgccaagc	acgtccccat	gcgctccatc	18540
aagaagagcg	acttcgcgga	gctggtgaag	tacatcaccg	acgagcaagg	caagaccgag	18600
cgcctttgcg	acgctca				·	18617

```
<210> 51
```

<213> Artificial

<220>

<223> Plasmid

<220>

<221> misc_feature

<222> (10264)..(10264)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10472)..(10472)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (10563)..(10563)

<223> n is a, c, g, or t

<211> 18333

<212> DNA

WO 2004/063359 PCT/EP2004/000099 306/357

<400> 51 ccgggctggt tgccctcgcc gctgggctgg cggccgtcta tggccctgca aacgcgccag 60 120 aaacgccgtc gaagccgtgt gcgagacacc gcggccgccg gcgttgtgga tacctcgcgg 180 aaaacttggc cctcactgac agatgagggg cggacgttga cacttgaggg gccgactcac 240 ccggcgcggc gttgacagat gaggggcagg ctcgatttcg gccggcgacg tggagctggc 300 cagcctcgca aatcggcgaa aacgcctgat tttacgcgag tttcccacag atgatgtgga 360 caagcctggg gataagtgcc ctgcggtatt gacacttgag gggcgcgact actgacagat 420 gaggggcgcg atccttgaca cttgaggggc agagtgctga cagatgaggg gcgcacctat tgacatttga ggggctgtcc acaggcagaa aatccagcat ttgcaagggt ttccgcccgt 480 540 ttttcggcca ccgctaacct gtcttttaac ctgcttttaa accaatattt ataaaccttg 600 tttttaacca gggctgcgcc ctgtgcgcgt gaccgcgcac gccgaagggg ggtgcccccc cttctcgaac cctcccggcc cgctaacgcg ggcctcccat ccccccaggg gctgcgcccc 660 720 teggeegega aeggeeteae eccaaaaatg geagegetgg eagteettge eattgeeggg atcggggcag taacgggatg ggcgatcagc ccgagcgcga cgcccggaag cattgacgtg 780 ccgcaggtgc tggcatcgac attcagcgac caggtgccgg gcagtgaggg cggcggcctg 840 ggtggcggcc tgcccttcac ttcggccgtc ggggcattca cggacttcat ggcggggccg 900 gcaattttta cettgggcat tettggcata gtggtegegg gtgeegtget egtgtteggg 960 ggtgcgataa acccagcgaa ccatttgagg tgataggtaa gattataccg aggtatgaaa 1020 acgagaattg gacctttaca gaattactct atgaagcgcc atatttaaaa agctaccaag 1080 acgaagagga tgaagaggat gaggaggcag attgccttga atatattgac aatactgata 1140 1200 agataatata tettttatat agaagatate geegtatgta aggattteag ggggeaagge ataggcagcg cgcttatcaa tatatctata gaatgggcaa agcataaaaa cttgcatgga 1260

WO 2004/063359 PCT/EP2004/000099 307/357

ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	taattgggta	1320
atgactccaa	cttattgata	gtgttttatg	ttcagataat	gcccgatgac	tttgtcatgc	1380
agctccaccg	attttgagaa	cgacagcgac	ttccgtccca	gccgtgccag	gtgctgcctc	1440
agattcaggt	tatgccgctc	aattegetge	gtatatcgct	tgctgattac	gtgcagcttt	1500
cccttcaggc	gggattcata	cageggeeag	ccatccgtca	tccatatcac	cacgtcaaag	1560
ggtgacagca	ggctcataag	acgccccagc	gtcgccatag	tgcgttcacc	gaatacgtgc	1620
gcaacaaccg	tcttccggag	actgtcatac	gcgtaaaaca	gccagcgctg	gcgcgattta	1680
gccccgacat	agccccactg	ttcgtccatt	tccgcgcaga	cgatgacgtc	actgcccggc	1740
tgtatgcgcg	aggttaccga	ctgcggcctg	agttttttaa	gtgacgtaaa	atcgtgttga	1800
ggccaacgcc	cataatgcgg	gctgttgccc	ggcatccaac	gccattcatg	gccatatcaa	1860
tgattttctg	gtgcgtaccg	ggttgagaag	cggtgtaagt	gaactgcagt	tgccatgttt	1920
tacggcagtg	agagcagaga	tagcgctgat	gtccggcggt	gcttttgccg	ttacgcacca	1980
ccccgtcagt	agctgaacag	gagggacagc	tgatagacac	agaagccact	ggagcacctc.	2040
aaaaacacca	tcatacacta	aatcagtaag	ttggcagcat	cacccataat	tgtggtttca	2100
aaatcggctc	cgtcgatact	atgttatacg	ccaactttga	aaacaacttt	gaaaaagctg	2160
ttttctggta	tttaaggttt	tagaatgcaa	ggaacagtga	attggagttc	gtcttgttat	2220
aattagcttc	ttggggtatc	tttaaatact	gtagaaaaga	ggaaggaaat	aataaatggc	2280
taaaatgaga	atatcaccgg	aattgaaaaa	actgatcgaa	aaataccgct	gcgtaaaaga	2340
tacggaagga	atgtctcctg	ctaaggtata	taagctggtg	ggagaaaatg	aaaacctata	2400
tttaaaaatg	acggacagcc	ggtataaagg	gaccacctat	gatgtggaac	gggaaaagga	2460
catgatgcta	tggctggaag	gaaagctgcc	tgttccaaag	gtcctgcact	ttgaacggca	2520

WO 2004/063359 PCT/EP2004/000099 308/357

tgatggctgg	agcaatctgc	tcatgagtga	ggccgatggc	gtcctttgct	cggaagagta	2580
tgaagatgaa	caaagccctg	aaaagattat	cgagctgtat	gcggagtgca	tcaggctctt	2640
tcactccatc	gacatatcgg	attgtcccta	tacgaatagc	ttagacagcc	gcttagccga	2700
attggattac	ttactgaata	acgatctggc	cgatgtggat	tgcgaaaact	gggaagaaga	2760
cactccattt	aaagateege	gcgagctgta	tgatttttta	aagacggaaa	agcccgaaga	2820
ggaacttgtc	ttttcccacg	gcgacctggg	agacagcaac	atctttgtga	aagatggcaa	2880
agtaagtggc	tttattgatc	ttgggagaag	cggcagggcg	gacaagtggt	atgacattgc	2940
cttctgcgtc	cggtcgatca	gggaggatat	cggggaagaa	cagtatgtcg	agctattttt	3000
tgacttactg	gggatcaagc	ctgattggga	gaaaataaaa	tattatattt	tactggatga	3060
attgttttag	tacctagatg	tggcgcaacg	atgccggcga	caagcaggag	cgcaccgact	3120
tcttccgcat	caagtgtttt	ggctctcagg	ccgaggccca	cggcaagtat	ttgggcaagg	3180
ggtcgctggt	attcgtgcag	ggcaagattc	ggaataccaa	gtacgagaag	gacggccaga	3240
cggtctacgg	gaccgacttc	attgccgata	aggtggatta	tctggacacc	aaggcaccag	3300
gcgggtcaaa	tcaggaataa	gggcacattg	cccggcgtg	agtcggggca	atcccgcaag	3360
gagggtgaat	gaatcggacg	tttgaccgga	aggcatacag	gcaagaactg	atcgacgcgg	3420
ggttttccgc	cgaggatgcc	gaaaccatcg	caageegeae	cgtcatgcgt	gegeeeegeg	3480
aaaccttcca	gteegtegge	tcgatggtcc	agcaagctac	ggccaagatc	gagcgcgaca	3540
gcgtgcaact	ggctccccct	gccctgcccg	cgccatcggc	cgccgtggag	cgttcgcgtc	3600
gtctcgaaca	ggaggcggca	ggtttggcga	agtcgatgac	catcgacacg	cgaggaacta	3660
tgacgaccaa	gaagcgaaaa	accgccggcg	aggacctggc	aaaacaggtc	agcgaggcca	3720
agcaggccgc	gttgctgaaa	cacacgaagc	agcagatcaa	ggaaatgcag	ctttccttgt	3780

WO 2004/063359 PCT/EP2004/000099 309/357

tcgatattgc	gccgtggccg	gacacgatgc	gagcgatgcc	aaacgacacg	gcccgctctg	3840
ccctgttcac	cacgcgcaac	aagaaaatcc	cgcgcgaggc	gctgcaaaac	aaggtcattt	3900
tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	gccgacgatg	3960
acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctatc	ggcgagccga	4020
tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	ggccggtatt	4080
acacgaaggc	cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	ttcacgtccg	4140
accgcgttgg	gcacctggaa	teggtgtege	tgctgcaccg	cttccgcgtc	ctggaccgtg	4200
gcaagaaaac	gtcccgttgc	caggtcctga	tcgacgagga	aatcgtcgtg	ctgtttgctg	4260
gcgaccacta	cacgaaattc	atatgggaga	agtaccgcaa	gctgtcgccg	acggcccgac	4320
ggatgttcga	ctatttcagc	tegeaceggg	agccgtaccc	gctcaagctg	gaaaccttcc	4380
gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	gtcggcgaag	4440
cctgcgaaga	gttgcgaggc	agcggcctgg	tggaacacgc	ctgggtcaat	gatgacctgg	4500
tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	gcagccagcg	4560
ctttactggc	atttcaggaa	caagegggea	ctgctcgacg	cacttgcttc	gctcagtatc	4620
gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	aattgacaat	4680
tgtgattaag	gctcagattc	gacggcttgg	ageggeegae	gtgcaggatt	tccgcgagat	4740
ccgattgtcg	gccctgaaga	aagctccaga	gatgttcggg	tccgtttacg	agcacgagga	4800
gaaaaagccc	atggaggcgt	tegetgaaeg	gttgcgagat	gccgtggcat	teggegeeta	4860
catcgacggo	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	ccaaggacgc	4920
tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagegaggee	gaggggtcgc	4980
cggtatgctg	g ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	, tccgacagat	5040

WO 2004/063359 PCT/EP2004/000099 310/357

tccaacggga atctggtgga tgcgcatctt catcctcggc gcacttaata tttcgctatt 5100 5160 ctggagettg ttgtttattt cggtctaccg cctgccgggc ggggtcgcgg cgacggtagg cgctgtgcag ccgctgatgg tcgtgttcat ctctgccgct ctgctaggta gcccgatacg 5220 5280 attgatggcg gtcctggggg ctatttgcgg aactgcgggc gtggcgctgt tggtgttgac accaaacgca gcgctagatc ctgtcggcgt cgcagcgggc ctggcggggg cggtttccat 5340 ggcgttcgga accgtgctga cccgcaagtg gcaacctccc gtgcctctgc tcacctttac 5400 cgcctggcaa ctggcggccg gaggacttct gctcgttcca gtagctttag tgtttgatcc 5460 5520 gccaatcccg atgcctacag gaaccaatgt tctcggcctg gcgtggctcg gcctgatcgg agcgggttta acctacttcc tttggttccg ggggatctcg cgactcgaac ctacagttgt 5580 ttccttactg ggctttctca gccccagatc tggggtcgat cagccgggga tgcatcaggc 5640 5700 cgacagtcgg aacttcgggt ccccgacctg taccattcgg tgagcaatgg ataggggagt 5760 tgatatcgtc aacgttcact tctaaagaaa tagcgccact cagcttcctc agcggcttta 5820 tccagcgatt tcctattatg tcggcatagt tctcaagatc gacagcctgt cacggttaag 5880 cgagaaatga ataagaaggc tgataattcg gatctctgcg agggagatga tatttgatca caggeageaa egetetgtea tegttacaat caacatgeta eeeteegega gateateegt 5940 gtttcaaacc cggcagctta gttgccgttc ttccgaatag catcggtaac atgagcaaag 6000 6060 totgoogcot tacaacggot otcocgotga cgccgtcccg gactgatggg ctgcctgtat 6120 cgagtggtga ttttgtgccg agctgccggt cggggagctg ttggctggct ggtggcagga tatattgtgg tgtaaacaaa ttgacgctta gacaacttaa taacacattg cggacgtttt 6180 taatgtactg gggtggtttt tcttttcacc agtgagacgg gcaacagctg attgcccttc 6240 6300 accgcctggc cctgagagag ttgcagcaag cggtccacgc tggtttgccc cagcaggcga

WO 2004/063359 PCT/EP2004/000099 311/357

aaatcctgtt	tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	aaagaatagc	6360
ccgagatagg	gttgagtgtt	gttccagttt	ggaacaagag	tccactatta	aagaacgtgg	6420
actccaacgt	caaagggcga	aaaaccgtct	atcagggcga	tggcccacta	cgtgaaccat	6480
cacccaaatc	aagttttttg	gggtcgaggt	gccgtaaagc	actaaatcgg	aaccctaaag	6540
ggagcccccg	atttagagct	tgacggggaa	agccggcgaa	cgtggcgaga	aaggaaggga	6600
agaaagcgaa	aggagcgggc	gccattcagg	ctgcgcaact	gttgggaagg	gcgatcggtg	6660
cgggcctctt	cgctattacg	ccagctggcg	aaagggggat	gtgctgcaag	gcgattaagt	6720
tgggtaacgc	cagggttttc	ccagtcacga	cgttgtaaaa	cgacggccag	tgaattcgag	6780
ctcggtaccc	ggggatcttt	cgacactgaa	atacgtcgag	cctgctccgc	ttggaagcgg	6840
cgaggagcct	cgtcctgtca	caactaccaa	catggagtac	gataagggcc	agttccgcca	6900
gctcattaag	agccagttca	tgggcgttgg	catgatggcc	gtcatgcatc	tgtacttcaa	6960
gtacaccaac	gctcttctga	tccagtcgat	cateegetga	aggcgctttc	gaatctggtt	7020
aagatccacg	tcttcgggaa	gccagcgact	ggtgacctcc	agcgtccctt	taaggctgcc	7080
aacagctttc	tcagccaggg	ccagcccaag	accgacaagg	cctccctcca	gaacgccgag	7140
aagaactgga	ggggtggtgt	caaggaggag	taagctcctt	attgaagtcg	gaggacggag	7200
cggtgtcaag	aggatattct	tcgactctgt	attatagata	agatgatgag	gaattggagg	7260
tagcatagct	tcatttggat	ttgctttcca	ggctgagact	ctagcttgga	gcatagaggg	7320
tcctttggct	ttcaatattc	tcaagtatct	cgagtttgaa	cttattccct	gtgaaccttt	7380
tattcaccaa	tgagcattgg	aatgaacatg	aatctgagga	ctgcaatcgc	catgaggttt	7440
tcgaaataca	tccggatgtc	gaaggcttgg	ggcacctgcg	ttggttgaat	ttagaacgtg	7500
gcactattga	tcatccgata	gctctgcaaa	gggcgttgca	caatgcaagt	caaacgttgc	7560

tagcagttcc	aggtggaatg	ttatgatgag	cattgtatta	aatcaggaga	tatagcatga	7620
tctctagtta	gctcaccaca	aaagtcagac	ggcgtaacca	aaagtcacac	aacacaagct	7680
gtaaggattt	cggcacggct	acggaagacg	gagaagccac	cttcagtgga	ctcgagtacc	7740
atttaattct	atttgtgttt	gatcgagacc	taatacagcc	cctacaacga	ccatcaaagt	7800
cgtatagcta	ccagtgagga	agtggactca	aatcgacttc	agcaacatct	cctggataaa	7860
ctttaagcct	aaactataca	gaataagata	ggtggagagc	ttataccgag	ctcccaaatc	7920
tgtccagatc	atggttgacc	ggtgcctgga	tcttcctata	gaatcatcct	tattcgttga	7980
cctagctgat	tctggagtga	cccagagggt	catgacttga	gcctaaaatc	cgccgcctcc	8040
accatttgta	gaaaaatgtg	acgaactcgt	gagctctgta	cagtgaccgg	tgactctttc	8100
tggcatgcgg	agagacggac	ggacgcagag	agaagggctg	agtaataagc	cactggccag	8160
acagetetgg	cggctctgag	gtgcagtgga	tgattattaa	tccgggaccg	gccgcccctc	8220
cgcccgaag	tggaaaggct	ggtgtgccc	: tcgttgacca	agaatctatt	gcatcatcgg	8280
agaatatgga	gcttcatcga	atcaccggca	gtaagcgaag	gagaatgtga	agccaggggt	8340
gtatagccgt	. cggcgaaata	gcatgccatt	: aacctaggta	cagaagtcca	attgcttccg	8400
atctggtaaa	a agattcacga	ı gatagtaccı	tctccgaagt	: aggtagagcg	agtacccggc	8460
gcgtaagcto	c cctaattggc	: ccatccggc	a tctgtagggo	gtccaaatat	cgtgcctctc	8520
ctgctttgc	c cggtgtatga	a aaccggaaa	g geegeteagg	g agctggccag	g cggcgcagac	8580
cgggaacac	a agctggcagt	cgacccatc	c ggtgctctgd	e actegacets	g ctgaggtccc	8640
tcagtccct	g gtaggcagc	t ttgccccgt	c tgtccgccc	g gtgtgtcggd	ggggttgaca	8700
aggtcgttg	c gtcagtcca	a catttgttg	c catattttc	c tgeteteec	c accagctgct	8760
cttttcttt	t ctctttctt	t toccatctt	c agtatattc	a tcttcccat	c caagaacctt	8820

WO 2004/063359 PCT/EP2004/000099 313/357

tatttcccct	aagtaagtac	tttgctacat	ccatactcca	tccttcccat	cccttattcc	8880
tttgaacctt	tcagttcgag	ctttcccact	tcatcgcagc	ttgactaaca	gctaccccgc	8940
ttgagcagac	atcaccatgc	ctgaactcac	cgcgacgtct	gtcgagaagt	ttctgatcga	9000
aaagttcgac	agcgtctccg	acctgatgca	gctctcggag	ggcgaagaat	ctcgtgcttt	9060
cagcttcgat	gtaggagggc	gtggatatgt	cctgcgggta	aatagctgcg	ccgatggttt	9120
ctacaaagat	cgttatgttt	atcggcactt	tgcatcggcc	gegeteeega	ttccggaagt	9180
gcttgacatt	ggggaattca	gcgagagcct	gacctattgc	atctcccgcc	gtgcacaggg	9240
tgtcacgttg	caagacctgc	ctgaaaccga	actgcccgct	gttctgcagc	cggtcgcgga	9300
ggccatggat	gcgatcgctg	cggccgatct	tagccagacg	agcgggttcg	gcccattcgg	9360
accgcaagga	atcggtcaat	acactacatg	gcgtgatttc	atatgcgcga	ttgctgatcc	9420
ccatgtgtat	cactggcaaa	ctgtgatgga	cgacaccgtc	agtgcgtccg	tegegeagge	9480
tctcgatgag	ctgatgcttt	gggccgagga	ctgccccgaa	gtccggcacc	tcgtgcacgc	9540
ggatttcggc	tccaacaatg	tcctgacgga	caatggccgc	ataacagcgg	tcattgactg	9600
gagcgaggcg	atgttcgggg	attcccaata	cgaggtcgcc	aacatcttct	tctggaggcc	9660
gtggttggct	tgtatggagc	agcagacgcg	ctacttcgag	cggaggcatc	cggagcttgc	9720
aggatcgccg	cggctccggg	cgtatatgct	ccgcattggt	cttgaccaac	tctatcagag	9780
cttggttgac	ggcaatttcg	atgatgcagc	ttgggcgcag	ggtcgatgcg	acgcaatcgt	9840
ccgatccgga	gccgggactg	tegggegtae	acaaatcgcc	cgcagaagcg	cggccgtctg	9900
gaccgatggc	tgtgtagaag	tactcgccga	tagtggaaac	cgacgcccca	gcactcgtcc	9960
gagggcaaag	gaatagagta	gatgccgacc	gcgggatcga	tccacttaac	gttactgaaa	10020
tcatcaaaca	gcttgacgaa	tctggatata	agategttgg	tgtcgatgtc	agctccggag	10080

WO 2004/063359 PCT/EP2004/000099 314/357

ttgagacaaa	tggtgttcag	gatctcgata	agatacgttc	atttgtccaa	gcagcaaaga	10140
gtgccttcta	gtgatttaat	agctccatgt	caacaagaat	aaaacgcgtt	ttcgggttta	10200
cctcttccag	atacagetea	tctgcaatgc	attaatgcat	tgactgcaac	ctagtaacgc	10260
cttncaggct	ccggcgaaga	gaagaatagc	ttagcagagc	tattttcatt	ttcgggagac	10320
gagatcaagc	agatcaacgg	tcgtcaagag	acctacgaga	ctgaggaatc	cgctcttggc	10380
tccacgcgac	tatatatttg	tctctaattg	tactttgaca	tgctcctctt	ctttactctg	10440
atagettgae	tatgaaaatt	ccgtcacdag	cncctgggtt	cgcaaagata	attgcatgtt	10500
tcttccttga	actctcaagc	ctacaggaca	cacattcatc	gtaggtataa	acctcgaaat	10560
canttcctac	taagatggta	tacaatagta	accatgcatg	gttgcctagt	gaatgctccg	10620
taacacccaa	tacgccggcc	gaaacttttt	tacaactctc	ctatgagtcg	tttacccaga	10680
atgcacaggt	acacttgttt	agaggtaatc	cttctttcta	gctagaagtc	ctcgtgtact	10740
gtgtaagege	ccactccaca	tctccactcg	acctgcaggc	atgcaagctt	gagattaaaa	10800
tagataagga	aaagaaagtg	aaaagaaatt	cggaagcatg	gcacattctt	ctttttataa	10860
atacatgcct	gactttcttt	ttccatcgat	atgatatatg	catatgatag	atatacaagc	10920
aatcttcttc	aaggagtttg	aaattttgtc	ctccaggagc	aaaaaaaagt	tttttttat	10980
acatgtttgt	acacaagaat	agttaccaat	ttgctttggt	cttacgtgct	gcaagtttat	11040
atcgttttca	atttctttgt	ctttacattt	tetttgteet	ttatctttcc	toatttagtc	11100
tttgggagaa	ttaggaaaag	ggagcggaaa	ggtaagaaat	gcttgcgtat	tttactaatt	11160
cggcaaacat	ccaatttggc	aaacagcagc	ctgtgcaacg	ctctcgagat	gacagtatct	11220
ttgattacac	tctaaatctc	gatgacccga	. ccaaaaagag	cgaacaaaga	aataatcttg	11280
tgcattcgaa	tatgatggaa	gatttttcc	cccttattct	aaatgttgac	atagcgtgta	11340

WO 2004/063359 PCT/EP2004/000099 315/357

tgttatataa	acaaaaagaa	attgtacaaa	ctttctttc	ttctcttttt	attttatctc	11400
tatgttgtgg	atttggaatg	ccctgatcgt	tttcgttacc	gtgattggca	tggaagtgat	11460
tgctgcactg	gcacacaaat	acatcatgca	cggctggggt	tggggatggc	atctttcaca	11520
tcatgaaccg	cgtaaaggtg	cgtttgaagt	taacgatctt	tatgccgtgg	tttttgctgc	11580
attatcgatc	ctgctgattt	atctgggcag	tacaggaatg	tggccgctcc	agtggattgg	11640
cgcaggtatg	acggcgtatg	gattactcta	ttttatggtg	cacgacgggc	tggtgcatca	11700
acgttggcca	ttccgctata	ttccacgcaa	gggctacctc	aaacggttgt	atatggcgca	11760
ccgtatgcat	cacgccgtca	ggggcaaaga	aggttgtgtt	tettttgget	tcctctatgc	11820
gccgcccctg	tcaaaacttc	aggcgacgct	ccgggaaaga	catggcgcta	gagcgggcgc	11880
tgccagagat	gegeagggeg	gggaggatga	gcccgcatcc	gggaagtaag	ggcctgacca	11940
gaggcggcca	gcagcagcgt	taatttttcg	ggcgtggtcg	ttgactgccg	ctgatcccaa	12000
agcttgagat	taaaatagat	aaggaaaaga	aagtgaaaag	aaattcggaa	gcatggcaca	12060
ttcttcttt	tataaataca	tgcctgactt	tctttttcca	tcgatatgat	atatgcatat	12120
gatagatata	caagcaatct	tcttcaagga	gtttgaaatt	ttgtcctcca	ggagcaaaaa	12180
aaagttttt	tttatacatg	tttgtacaca	agaatagtta	ccaatttgct	ttggtcttac	12240
gtgctgcaag	tttatatcgt	tttcaatttc	tttgtcttta	cattttcttt	gtcctttatc	12300
tttcctcatt	tagtctttgg	gagaattagg	aaaagggagc	ggaaaggtaa	gaaatgcttg	12360
cgtattttac	taattcggca	aacatccaat	ttggcaaaca	gcagcctgtg	caacgctctc	12420
gagatgacag	tatctttgat	tacactctaa	atctcgatga	cccgaccaaa	aagagcgaac	12480
aaagaaataa	tcttgtgcat	tcgaatatga	tggaagattt	tttccccctt	attctaaatg	12540
ttgacatago	gtgtatgtta	tataaacaaa	aagaaattgt	acaaactttc	ttttcttctc	12600

WO 2004/063359 PCT/EP2004/000099 316/357

ttttattt	atctctatga	tccagttaga	acaaccactc	agtcatcaag	caaaactgac	12660
tccagtactg	agaagtaaat	ctcagtttaa	ggggcttttc	attgctattg	tcattgttag	12720
cgcatgggtc	attagcctga	gtttattact	ttcccttgac	atctcaaagc	taaaattttg	12780
gatgttattg	cctgttatac	tatggcaaac	atttttatat	acgggattat	ttattacatc	12840
tcatgatgcc	atgcatggcg	tagtatttcc	ccaaaacacc	aagattaatc	atttgattgg	12900
aacattgacc	ctatcccttt	atggtctttt	accatatcaa	aaactattga	aaaaacattg	12960
gttacaccac	cacaatccag	caagctcaat	agacccggat	tttcacaatg	gtaaacacca	13020
aagtttcttt	gcttggtatt	ttcattttat	gaaaggttac	tggagttggg	ggcaaataat	13080
tgcgttgact	attatttata	actttgctaa	atacatactc	catateceaa	gtgataatct	13140
aacttacttt	tgggtgctac	cctcgctttt	aagttcatta	caattattct	attttggtac	13200
tittttaccc	catagtgaac	caataggggg	ttatgttcag	cctcattgtg	cccaaacaat	13260
tagccgtcct	atttggtggt	catttatcac	gtgctatcat	tttggctacc	acgaggaaca	13320
tcacgaatat	cctcatattt	cttggtggca	gttaccagaa	atttacaaag	caaaatagaa	13380
gcttggcgta	atcatggtca	tagctgtttc	ctgtgtgaaa	ttgttatccg	ctcacaattc	13440
cacacaacat	acgagccgga	agcataaagt	gtaaagcctg	gggtgcctaa	tgagtgagct	13500
aactcacatt	aattgcgttg	cgctcactgc	ccgctttcca	gtcgggaaac	ctgtcgtgcc	13560
agctgcatta	atgaatcggc	caacgcgcgg	ggagaggcgg	tttgcgtatt	gggccaaaga	13620
caaaagggcg	acattcaacc	gattgaggga	gggaaggtaa	atattgacgg	aaattattca	13680
ttaaaggtga	attatcaccg	tcaccgactt	gagccatttg	ggaattagag	ccagcaaaat	13740
caccagtage	accattacca	ttagcaaggc	cggaaacgtc	accaatgaaa	ccatcgatag	13800
cagcacegta	atcagtagcg	acagaatcaa	gtttgccttt	agcgtcagac	tgtagcgcgt	13860

WO 2004/063359 PCT/EP2004/000099 317/357

tttcatcggc	attttcggtc	atagececet	tattagcgtt	tgccatcttt	tcataatcaa	13920
aatcaccgga	accagagcca	ccaccggaac	cgcctccctc	agageegeea	ccctcagaac	13980
cgccaccctc	agagccacca	ccctcagagc	cgccaccaga	accaccacca	gagccgccgc	14040
cagcattgac	aggaggcccg	atctagtaac	atagatgaca	ccgcgcgcga	taatttatcc	14100
tagtttgcgc	gctatatttt	gttttctatc	gcgtattaaa	tgtataattg	cgggactcta	14160
atcataaaaa	cccatctcat	aaataacgtc	atgcattaca	tgttaattat	tacatgctta	14220
acgtaattca	acagaaatta	tatgataatc	atcgcaagac	cggcaacagg	attcaatctt	14280
aagaaacttt	attgccaaat	gtttgaacga	tcggggatca	tccgggtctg	tggcgggaac	14340
tccacgaaaa	tatccgaacg	cagcaagata	tcgcggtgca	tctcggtctt	gcctgggcag	14400
tegeegeega	cgccgttgat	gtggacgccg	ggcccgatca	tattgtcgct	caggatcgtg	14460
gcgttgtgct	tgtcggccgt	tgctgtcgta	atgatatcgg	caccttcgac	cgcctgttcc	14520
gcagagatcc	cgtgggcgaa	gaactccagc	atgagatccc	cgcgctggag	gatcatccag	14580
ccggcgtccc	ggaaaacgat	tccgaagccc	aacctttcat	agaaggcggc	ggtggaatcg	14640
aaatctcgtg	atggcaggtt	gggcgtcgct	tggtcggtca	tttcgaaccc	cagagtcccg	14700
ctcagaagaa	ctcgtcaaga	aggcgataga	aggcgatgcg	ctgcgaatcg	ggagcggcga	14760
taccgtaaag	cacgaggaag	cggtcagccc	attcgccgcc	aagctcttca	gcaatatcac	14820
gggtagccaa	cgctatgtcc	tgatagcggt	ccgccacacc	cageeggeea	cagtcgatga	14880
atccagaaaa	gcggccattt	tccaccatga	tattcggcaa	gcaggcatcg	ccatgggtca	14940
cgacgagatc	atcgccgtcg	ggcatgcgcg	ccttgagcct	ggcgaacagt	teggetggeg	15000
cgagcccctg	atgctcttcg	tccagatcat	cctgatcgac	aagaccggct	tccatccgag	15060
tacgtgctcg	ctcgatgcga	tgtttcgctt	ggtggtcgaa	tgggcaggta	gccggatcaa	15120

WO 2004/063359 PCT/EP2004/000099 318/357

gcgtatgcag	ccgccgcatt	gcatcagcca	tgatggatac	tttctcggca	ggagcaaggt	15180
gagatgacag	gagateetge	cccggcactt	cgcccaatag	cagccagtcc	cttcccgctt	15240
cagtgacaac	gtcgagcaca	gctgcgcaag	gaacgcccgt ·	cgtggccagc	cacgatagcc	15300
gegetgeete	gtcctgcagt	tcattcaggg	caccggacag	gtcggtcttg	acaaaaagaa	15360
ccgggcgccc	ctgcgctgac	agccggaaca	cggcggcatc	agagcagccg	attgtctgtt	15420
gtgcccagtc	atagccgaat	agcctctcca	cccaagcggc	cggagaacct	gcgtgcaatc	15480
catcttgttc	aatcatgcga	aacgatccag	atccggtgca	gattatttgg	attgagagtg	15540
aatatgagac	tctaattgga	taccgagggg	aatttatgga	acgtcagtgg	agcatttttg	15600
acaagaaata	tttgctagct	gatagtgacc	ttaggcgact	tttgaacgcg	caataatggt	15660
ttctgacgta	tgtgcttagc	tcattaaact	ccagaaaccc	gcggctgagt	ggctccttca	15720
acgttgcggt	tctgtcagtt	ccaaacgtaa	aacggcttgt	cccgcgtcat	cggcgggggt	15780
cataacgtga	ctcccttaat	teteegetea	tgatcagatt	gtcgtttccc	gccttcagtt	15840
taaactatca	gtgtttgaca	ggatatattg	gcgggtaaac	ctaagagaaa	agagcgttta	15900
ttagaataat	cggatattta	aaagggcgtg	aaaaggttta	teegttegte	catttgtatg	15960
tgcatgccaa	ccacagggtt	ccccagatct	ggcgccggcc	agcgagacga	gcaagattgg	16020
ccgccgcccg	aaacgatccg	acagegegee	cagcacaggt	gcgcaggcaa	attgcaccaa	16080
cgcatacagc	gccagcagaa	tgccatagtg	ggcggtgacg	tcgttcgagt	gaaccagatc	16140
gcgcaggagg	cccggcagca	ccggcataat	caggccgatg	ccgacagcgt	cgagcgcgac	16200
agtgctcaga	attacgatca	ggggtatgtt	gggtttcacg	tetggeetee	ggaccagcct	16260
cegetgątee	gattgaacgc	gcggattctt	tatcactgat	aagttggtgg	acatattatg	16320
tttatcagtg	ataaagtgtc	aagcatgaca	aagttgcagc	cgaatacagt	gatccgtgcc	16380

WO 2004/063359 PCT/EP2004/000099 319/357

gccctggacc	tgttgaacga	ggtcggcgta	gacggtctga	cgacacgcaa	actggcggaa	16440
cggttggggg	ttcagcagcc	ggcgctttac	tggcacttca	ggaacaagcg	ggcgctgctc	16500
gacgcactgg	ccgaagccat	gctggcggag	aațcatacgc	attcggtgcc	gagageegae	16560
gacgactggc	gctcatttct	gatcgggaat	gcccgcagct	tcaggcaggc	gctgctcgcc	16620
taccgcgatg	gcgcgcgcat	ccatgccggc	acgcgaccgg	gcgcaccgca	gatggaaacg	16680
gccgacgcgc	agcttcgctt	cctctgcgag	gcgggttttt	cggccgggga	cgccgtcaat	16740
gcgctgatga	caatcagcta	cttcactgtt	ggggccgtgc	ttgaggagca	ggccggcgac	16800
agcgatgccg	gcgagcgcgg	cggcaccgtt	gaacaggctc	cgctctcgcc	gctgttgcgg	16860
gccgcgatag	acgccttcga	cgaagccggt	ccggacgcag	cgttcgagca	gggactcgcg	16920
gtgattgtcg	atggattggc	gaaaaggagg	ctcgttgtca	ggaacgttga	aggaccgaga	16980
aagggtgacg	attgatcagg	accgctgccg	gagegeaace	cactcactac	agcagagcca	17040
tgtagacaac	atcccctccc	cctttccacc	gcgtcagacg	cccgtagcag	cccgctacgg	17100
gctttttcat	gccctgccct	agcgtccaag	cctcacggcc	gegeteggee	tctctggcgg	17160
ccttctggcg	ctcttccgct	tcctcgctca	ctgactcgct	gegeteggte	gttcggctgc	17220
ggcgagcggt	atcagctcac	tcaaaggcgg	taatacggtt	atccacagaa	tcaggggata	17280
acgcaggaaa	gaacatgtga	gcaaaaggcc	agcaaaaggc	caggaaccgt	aaaaaggccg	17340
cgttgctggc	gtttttccat	aggctccgcc	.cccctgacga	gcatcacaaa	aatcgacgct	17400
caagtcagag	gtggcgaaac	ccgacaggac	tataaagata	ccaggcgttt	ccccctggaa	17460
gctccctcgt	gegeteteet	gttccgaccc	tgccgcttac	cggatacctg	tccgcctttc	17520
tcccttcggg	aagcgtggcg	cttttccgct	gcataaccct	gcttcggggt	cattatagcg	17580
attttttcgg	tatatccatc	ctttttcgca	cgatatacag	gattttgcca	aagggttcgt	17640

gtagactttc cttggtgtat	ccaacggcgt	cageegggea	ggataggtga	agtaggccca	17700
cccgcgagcg ggtgttcctt	cttcactgtc	ccttattcgc	acctggcggt	gctcaacggg	17760
aatcctgctc tgcgaggctg	gccggctacc	gccggcgtaa	cagatgaggg	caagcggatg	17820
gctgatgaaa ccaagccaac	caggaagggc	agcccaccta	tcaaggtgta	ctgccttcca	17880
gacgaacgaa gagcgattga	ggaaaaggcg	geggeggeeg	gcatgagcct	gtcggcctac	17940
ctgctggccg tcggccaggg	ctacaaaatc	acgggcgtcg	tggactatga	gcacgtccgc	18000
gagetggeee geateaatgg	cgacctgggc	cgcctgggcg	gcctgctgaa	actctggctc	18060
accgacgacc cgcgcacggc	geggtteggt	gatgccacga	tcctcgccct	gctggcgaag	18120
atcgaagaga agcaggacga	gcttggcaag	gtcatgatgg	gcgtggtccg	cccgagggca	18180
gagccatgac ttttttagcc	gctaaaacgg	ccggggggtg	cgcgtgattg	ccaagcacgt	18240
ccccatgcgc tccatcaaga	a agagcgactt	cgcggagctg	gtgaagtaca	tcaccgacga	18300
gcaaggcaag accgagcgc	: tttgcgacgc	tca			18333

```
<210> 52
```

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<211> 17

<212> DNA

<213> Artificial

321/357

<222> (9)..(9)

<223> n is a, c, g, or t

<400> 52

17 gcngarggna thtggta

<210> 53

<211> 20

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3)..(3)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (6)..(6)

<223> n is a, c, g, or t

<400> 53

20 tengenagra adatrttrtg

27

<210> 54

<211> 27

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 54

aagtgacacc ggttacacgc ttgtctt

•	WO 2004/063359		322/357	PCT/EP2004/000099
<210>	55			
<211>	27			
<212>.	DNA			
<213>	Artificial			
<220>				
<223>	Primer			
<400>	55			
gcttat	cacc atctgttacc to	cttgc		27
<210>				
<211>	32			
	DNA			•
<213>	Artificial			
<220>				·
<223>	Primer			
<400>				
agagag	ggat ccttaaatgc ga	atatcgtt gc		32
<210>				
<211>	32			
<212>				
<213>	Artificial			
<220>				
	Postorona			
<223>	Primer			
<400>	57			
		2222222		32
ayayay	ggat ccatgtctga tc	aaaayaay ca		52

<211> 37 <212> DNA <213> Artificial

<210> 58

<220>

38

<220>

<223> Primer

ctattttaat catatgtctg atcaaaagaa gcatattg

<400> 61

WO 2004/063359 PCT/EP2004/000099

<210> 62

<211> 16103

<212> DNA

<213> Artificial

<220>

<223> Primer

<220>

<221> misc_feature

<222> (3471)..(3471)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3679)..(3679)

<223> n is a, c, g, or t

<220>

<221> misc_feature

<222> (3770)..(3770)

<223> n is a, c, g, or t

<400> 62

gatctttcga cactgaaata cgtcgagcct gctccgcttg gaagcggcga ggagcctcgt 60 cctgtcacaa ctaccaacat ggagtacgat aagggccagt tccgccagct cattaagagc 120 180 cagttcatgg gcgttggcat gatggccgtc atgcatctgt acttcaagta caccaacgct cttctgatcc agtcgatcat ccgctgaagg cgctttcgaa tctggttaag atccacgtct 240 tegggaagee agegactggt gacetecage gteeetttaa ggetgeeaac agetttetea 300 gccagggcca gcccaagacc gacaaggcct ccctccagaa cgccgagaag aactggaggg 360 420 gtggtgtcaa ggaggagtaa gctccttatt gaagtcggag gacggagcgg tgtcaagagg 480 atattcttcg actctgtatt atagataaga tgatgaggaa ttggaggtag catagcttca tttggatttg ctttccaggc tgagactcta gcttggagca tagagggtcc tttggctttc 540

WO 2004/063359 PCT/EP2004/000099 325/357

aatattctca	agtatctcga	gtttgaactt	attccctgtg	aaccttttat	tcaccaatga	600
gcattggaat	gaacatgaat	ctgaggactg	caatcgccat	gaggttttcg	aaatacatcc	660
ggatgtcgaa	ggcttggggc	acctgcgttg	gttgaattta	gaacgtggca	ctattgatca	720
tccgatagct	ctgcaaaggg	cgttgcacaa	tgcaagtcaa	acgttgctag	cagttccagg	780
tggaatgtta	tgatgagcat	tgtattaaat	caggagatat	agcatgatct	ctagttagct	840
caccacaaaa	gtcagacggc	gtaaccaaaa	gtcacacaac	acaagctgta	aggatttcgg	900
cacggctacg	gaagacggag	aagccacctt	cagtggactc	gagtaccatt	taattctatt	960
tgtgtttgat	cgagacctaa	tacagcccct	acaacgacca	tcaaagtcgt	atagctacca	1020
gtgaggaagt	ggactcaaat	cgacttcagc	aacatctcct	ggataaactt	taagcctaaa	1080
ctatacagaa	taagataggt	ggagagctta	taccgagctc	ccaaatctgt	ccagatcatg	1140
gttgaccggt	gcctggatct	tcctatagaa	tcatccttat	tegttgacet	agctgattct	1200
ggagtgaccc	agagggtcat	gacttgagcc	taaaatccgc	cgcctccacc	atttgtagaa	1260
aaatgtgacg	aactcgtgag	ctctgtacag	tgaccggtga	ctctttctgg	catgeggaga	1320
gacggacgga	cgcagagaga	agggctgagt	aataagccac	tggccagaca	gctctggcgg	1380
ctctgaggtg	cagtggatga	ttattaatcc	gggaccggcc	gcccctccgc	cccgaagtgg	1440
aaaggctggt	gtgcccctcg	ttgaccaaga	atctattgca	tcatcggaga	atatggagct	1500
tcatcgaatc	accggcagta	agcgaaggag	aatgtgaagc	caggggtgta	tagccgtcgg	1560
cgaaatagca	tgccattaac	ctaggtacag	aagtccaatt	gcttccgatc	tggtaaaaga	1620
ttcacgagat	agtaccttct	ccgaagtagg	tagagcgagt	acccggcgcg	taagctccct	1680
aattggccca	teeggeatet	gtagggcgtc	caaatatcgt	gcctctcctg	ctttgcccgg	1740
tgtatgaaac	: cggaaaggcc	gctcaggagc	tggccagcgg	cgcagaccgg	gaacacaagc	1800

WO 2004/063359 PCT/EP2004/000099 326/357

tggcagtcga	cccatccggt	gctctgcact	cgacctgctg	aggtccctca	gtccctggta	1860
ggcagctttg	ccccgtctgt	ccgcccggtg	tgtcggcggg	gttgacaagg	tcgttgcgtc	1920
agtccaacat	ttgttgccat	attttcctgc	tctccccacc	agctgctctt	ttettttete	1980
tttctttcc	catcttcagt	atattcatct	tcccatccaa	gaacctttat	ttcccctaag	2040
taagtacttt	gctacatcca	tactccatcc	ttcccatccc	ttattccttt	gaacctttca	2100
gttcgagctt	tcccacttca	tegeagettg	actaacagct	accccgcttg	agcagacatc	2160
accatgcctg	aactcaccgc	gacgtctgtc	gagaagtttc	tgatcgaaaa	gttcgacagc	2220
gtctccgacc	tgatgcagct	ctcggagggc	gaagaatctc	gtgctttcag	cttcgatgta	2280
ggagggcgtg	gatatgtcct	gcgggtaaat	agctgcgccg	atggtttcta	caaagatcgt	2340
tatgtttatc	ggcactttgc	atcggccgcg	ctcccgattc	cggaagtgct	tgacattggg	2400
gaattcagcg	agagcctgac	ctattgcatc	tecegeegtg	cacagggtgt	cacgttgcaa	2460
gacctgcctg	aaaccgaact	gcccgctgtt	ctgcagccgg	tegeggagge	catggatgcg	2520
atcgctgcgg	ccgatcttag	ccagacgagc	gggttcggcc	cattcggacc	gcaaggaatc	2580
ggtcaataca	ctacatggcg	tgatttcata	tgcgcgattg	ctgatcccca	tgtgtatcac	2640
tggcaaactg	tgatggacga	caccgtcagt	gcgtccgtcg	cgcaggctct	cgatgagctg	2700
atgctttggg	ccgaggactg	ccccgaagtc	cggcacctcg	tgcacgcgga	ttteggetee	2760
aacaatgtcc	tgacggacaa	tggccgcata	acageggtca	ttgactggag	cgaggcgatg	2820
ttcggggatt	cccaatacga	ggtcgccaac	atcttcttct	ggaggccgtg	gttggcttgt	2880
atggagcagc	agacgcgcta	cttcgagcgg	aggcatccgg	agcttgcagg	atcgccgcgg	2940
ctccgggcgt	atatgctccg	cattggtctt	gaccaactct	atcagagett	ggttgacggc	3000
aatttcgatg	atgcagcttg	ggcgcagggt	cgatgcgacg	caatcgtccg	atccggagcc	3060

WO 2004/063359 PCT/EP2004/000099 327/357

gggactgtcg	ggcgtacaca	aatcgcccgc	agaagcgcgg	ccgtctggac	cgatggctgt	3120
gtagaagtac	tcgccgatag	tggaaaccga	cgccccagca	ctcgtccgag	ggcaaaggaa	3180
tagagtagat	gccgaccgcg	ggatcgatcc	acttaacgtt	actgaaatca	tcaaacagct	3240
tgacgaatct	ggatataaga	tcgttggtgt	cgatgtcagc	tccggagttg	agacaaatgg	3300
tgttcaggat	ctcgataaga	tacgttcatt	tgtccaagca	gcaaagagtg	ccttctagtg	3360
atttaatagc	tccatgtcaa	caagaataaa	acgcgttttc	gggtttacct	cttccagata	3420
cageteatet	gcaatgcatt	aatgcattga	ctgcaaccta	gtaacgcctt	ncaggeteeg	3480
gcgaāgagaa	gaatagctta	gcagagctat	tttcattttc	gggagacgag	atcaagcaga	3540
tcaacggtcg	tcaagagacc	tacgagactg	aggaatccgc	tcttggctcc	acgcgactat	3600
atatttgtct	ctaattgtac	tttgacatgc	tectettett	tactctgata	gcttgactat	3660
gaaaattccg	tcaccagcnc	ctgggttcgc	aaagataatt	gcatgtttct	tccttgaact	3720
ctcaagccta	caggacacac	attcatcgta	ggtataaacc	tcgaaatcan	ttcctactaa	3780 ·
gatggtatac	aatagtaacc	atgcatggtt	gcctagtgaa	tgctccgtaa	cacccaatac	3840
gccggccgaa	acttttttac	aactctccta	tgagtcgttt	acccagaatg	cacaggtaca	3900
cttgtttaga	ggtaatcctt	ctttctagct	agaagtcctc	gtgtactgtg	taagcgccca	3960
ctccacatct	ccactcgacc	tgcaggcatg	caagcttgag	tctatcgcct	ccaaaaagta	4020
cggtgctgaa	ttcagatatc	aatcgcctgt	tgctaaaatt	aacactgtcg	ataaagacaa	4080
gcgtgtaacc	ggtgtcactt	tggaaagcgg	agaagtcatt	gaagccgatg	cagtcgtatg	4140
taatgcggat	cttgtttatg	cttatcacca	tctgttacct	ccttgcaatt	ggacaaagaa	4200
gacattagcc	tcaaagaaac	tcacttcatc	atctatttcg	ttttattggt	ccatgtcaac	4260
aaaggtgcct	caattagacg	tacacaatat	cttcttggct	gaagcctaca	aggaaagttt	4320

WO 2004/063359 PCT/EP2004/000099 328/357

tgatgagatt	ttcaacgact	tcggtttgcc	ctctgaagct	tggcgtaatc	atggtcatag	4380
ctgtttcctg	tgtgaaattg	ttatccgctc	acaattccac	acaacatacg	agccggaagc	4440
ataaagtgta	aagcctgggg	tgcctaatga	gtgagctaac	tcacattaat	tgcgttgcgc	4500
tcactgcccg	ctttccagtc	gggaaacctg	tegtgecage	tgcattaatg	aatcggccaa	4560
cgcgcgggga	gaggcggttt	gcgtattggg	ccaaagacaa	aagggcgaca	ttcaaccgat	4620
tgagggaggg	aaggtaaata	ttgacggaaa	ttattcatta	aaggtgaatt	atcaccgtca	4680
ccgacttgag	ccatttggga	attagagcca	gcaaaatcac	cagtagcacc	attaccatta	4740
gcaaggccgg	aaacgtcacc	aatgaaacca	tcgatagcag	caccgtaatc	agtagcgaca	4800
gaatcaagtt	tgcctttagc	gtcagactgt	agcgcgtttt	catcggcatt	ttcggtcata	4860
gccccttat	tagcgtttgc	catcttttca	taatcaaaat	caccggaacc	agagccacca	4920
ccggaaccgc	ctccctcaga	geegeeaeee	tcagaaccgc	caccctcaga	gccaccaccc	4980
tcagagcegc	caccagaacc	accaccagag	ccgccgccag	cattgacagg	aggcccgatc	5040
tagtaacata	gatgacaccg	cgcgcgataa	tttatcctag	tttgcgcgct	atattttgtt	5100
ttctatcgcg	tattaaatgt	ataattgcgg	gactctaatc	ataaaaaccc	atctcataaa	5160
taacgtcatg	cattacatgt	taattattac	atgcttaacg	taattcaaca	gaaattatat	5220
gataatcatc	gcaagaccgg	caacaggatt	caatcttaag	aaactttatt	gccaaatgtt	5280
tgaacgatcg	gggatcatcc	gggtctgtgg	cgggaactcc	acgaaaatat	ccgaacgcag	5340
caagatatcg	cggtgcatct	cggtcttgcc	tgggcagtcg	ccgccgacgc	cgttgatgtg	5400
gacgccgggc	ccgatcatat	tgtcgctcag	gatcgtggcg	ttgtgcttgt	cggccgttgc	5460
tgtcgtaatg	atateggeae	cttcgaccgc	ctgttccgca	gagatecegt	gggcgaagaa	5520
ctccagcatg	agateceege	gctggaggat	catccagccg	gcgtcccgga	aaacgattcc	5580

WO 2004/063359 PCT/EP2004/000099 329/357

gaagcccaac	ctttcataga	aggcggcggt	ggaatcgaaa	tctcgtgatg	gcaggttggg	5640
cgtcgcttgg	tcggtcattt	cgaaccccag	agtcccgctc	agaagaactc	gtcaagaagg	5700
cgatagaagg	cgatgcgctg	cgaatcggga	gcggcgatac	cgtaaagcac	gaggaagcgg	5760
tcagcccatt	cgccgccaag	ctcttcagca	atatcacggg	tagccaacgc	tatgtcctga	5820
tagcggtccg	ccacacccag	ccggccacag	tcgatgaatc	cagaaaagcg	gccattttcc	5880
accatgatat	tcggcaagca	ggcatcgcca	tgggtcacga	cgagatcatc	gcegteggge	5940
atgcgcgcct	tgagcctggc	gaacagttcg	gctggcgcga	gcccctgatg	ctcttcgtcc	6000
agatcatcct	gatcgacaag	accggcttcc	atccgagtac	gtgctcgctc	gatgcgatgt	6060
ttcgcttggt	ggtcgaatgg	gcaggtagcc	ggatcaagcg	tatgcagccg	ccgcattgca	6120
tcagccatga	tggatacttt	ctcggcagga	gcaaggtgag	atgacaggag	atcctgcccc	6180
ggcacttcgc	ccaatagcag	ccagtccctt	cccgcttcag	tgacaacgtc	gagcacagct	6240
gcgcaaggaa	cgcccgtcgt	ggccagccac	gatageegeg	ctgcctcgtc	ctgcagttca	6300
ttcagggcac	cggacaggtc	ggtcttgaca	aaaagaaccg	ggcgcccctg	cgctgacagc	6360
cggaacacgg	cggcatcaga	gcagccgatt	gtctgttgtg	cccagtcata	gccgaatagc	6420
ctctccaccc	aagcggccgg	agaacctgcg	tgcaatccat	cttgttcaat	catgcgaaac	. 6480
gatccagatc	cggtgcagat	tatttggatt	gagagtgaat	atgagactct	aattggatac	6540
cgaggggaat	ttatggaacg	tcagtggagc	atttttgaca	agaaatattt	gctagctgat	6600
agtgacctta	ggcgactttt	gaacgcgcaa	taatggtttc	tgacgtatgt	gcttagctca	6660
ttaaactcca	gaaacccgcg	gctgagtggc	tccttcaacg	ttgcggttct	gtcagttcca	6720
aacgtaaaac	ggcttgtccc	gcgtcatcgg	cgggggtcat	aacgtgactc	ccttaattct	6780
ccgctcatga	tcagattgtc	gtttcccgcc	ttcagtttaa	actatcagtg	tttgacagga	6840

WO 2004/063359 PCT/EP2004/000099 330/357

tatattggcg	ggtaaaccta	agagaaaaga	gcgtttatta	gaataatcgg	atatttaaaa	6900
gggcgtgaaa	aggtttatcc	gttcgtccat	ttgtatgtgc	atgccaacca	cagggttccc	6960
cagatctggc	gccggccagc	gagacgagca	agattggccg	ccgcccgaaa	cgatccgaca	7020
gcgcgcccag	cacaggtgcg	caggcaaatt	gcaccaacgc	atacagcgcc	agcagaatgc	7080
catagtgggc	ggtgacgtcg	ttcgagtgaa	ccagatcgcg	caggaggccc	ggcagcaccg	7140
gcataatcag	gccgatgccg	acagcgtcga	gcgcgacagt	gctcagaatt	acgatcaggg	7200
gtatgttggg	tttcacgtct	ggcctccgga	ccagcctccg	ctggtccgat	tgaacgcgcg	7260 ·
gattctttat	cactgataag	ttggtggaca	tattatgttt	atcagtgata	aagtgtcaag	7320
catgacaaag	ttgcagccga	atacagtgat	ccgtgccgcc	ctggacctgt	tgaacgaggt	7380
cggcgtagac	ggtctgacga	cacgcaaact	ggcggaacgg	ttgggggttc	agcagccggc	7440
gctttactgg	cacttcagga	acaagcgggc	gctgctcgac	gcactggccg	aagccatgct	7500
ggcggagaat	catacgcatt	cggtgccgag	agccgacgac	gactggcgct	catttctgat	7560
cgggaatgcc	cgcagcttca	ggcaggcgct	gctcgcctac	cgcgatggcg	cgcgcatcca	7620
tgccggcacg	cgaccgggcg	caccgcagat	ggaaacggcc	gacgcgcagc	ttegetteet	7680
ctgcgaggcg	ggtttttcgg	ccggggacgc	cgtcaatgcg	ctgatgacaa	tcagctactt	7740
cactgttggg	gccgtgcttg	aggagcaggc	cggcgacagc	gatgccggcg	agegeggegg	7800
caccgttgaa	caggeteege	tetegeegat	gttgcgggcc	gcgatagacg	ccttcgacga	7860
agccggtccg	gacgcagcgt	tcgagcaggg	actcgcggtg	attgtcgatg	gattggcgaa	7920
aaggaggctc	gttgtcagga	acgttgaagg	accgagaaag	ggtgacgatt	gatcaggacc	7980
gctgccggag	cgcaacccac	tcactacagc	agagccatgt	agacaacatc	ccctccccct	8040
ttccaccgcg	tcagacgccc	gtagcagccc	gctacgggct	ttttcatgcc	ctgccctagc	8100 .

WO 2004/063359 PCT/EP2004/000099 331/357

gtccaagcct	caeggeegeg	cteggeetet	ctggcggcct	tetggegete	ttccgcttcc	8160
tcgctcactg	actcgctgcg	ctcggtcgtt	cggctgcggc	gagcggtatc	agctcactca	8220
aaggcggtaa	tacggttatc	cacagaatca	ggggataacg	caggaaagaa	catgtgagca	8280
aaaggccagc	aaaaggccag	gaaccgtaaa	aaggccgcgt	tgctggcgtt	tttccatagg	8340
ctccgccccc	ctgacgagca	tcacaaaaat	cgacgctcaa	gtcagaggtg	gcgaaacccg	8400
acaggactat	aaagatacca	ggcgtttccc	cctggaagct	ccctcgtgcg	ctctcctgtt	8460
ccgaccctgc	cgcttaccgg	atacctgtcc	gcctttctcc	cttcgggaag	cgtggcgctt	8520
ttccgctgca	taaccctgct	tcggggtcat	tatagcgatt	ttttcggtat	atccatcctt	8580
tttcgcacga	tatacaggat	tttgccaaag	ggttcgtgta	gactttcctt	ggtgtatcca	8640
acggcgtcag	ccgggcagga	taggtgaagt	aggcccaccc	gcgagcgggt	gttccttctt	8700
cactgtccct	tattcgcacc	tggcggtgct	caacgggaat	cctgctctgc	gaggctggcc	8760
ggctaccgcc	ggcgtaacag	atgagggcaa	gcggatggct	gatgaaacca	agccaaccag	8820
gaagggcagc	ccacctatca	aggtgtactg	ccttccagac	gaacgaagag	cgattgagga	8880
aaaggcggcg	gcggccggca	tgagcctgtc	ggcctacctg	ctggccgtcg	gccagggcta	8940
caaaatcacg	ggcgtcgtgg	actatgagca	cgtccgcgag	ctggcccgca	tcaatggcga	9000
cctgggccgc	ctgggcggcc	tgctgaaact	ctggctcacc	gacgacccgc	gcacggcgcg	9060
gttcggtgat	gccacgatcc	tcgccctgct	ggcgaagatc	gaagagaagc	aggacgagct	9120
tggcaaggtc	atgatgggcg	tggtccgccc	gagggcagag	ccatgacttt	tttagccgct	9180
aaaacggccg	gggggtgcgc	gtgattgcca	agcacgtccc	catgcgctcc	atcaagaaga	9240
gcgacttcgc	ggagctggtg	aagtacatca	ccgacgagca	aggcaagacc	gagcgccttt	9300
gcgacgctca	ccgggctggt	tgccctcgcc	gctgggctgg	cggccgtcta	tggccctgca	9360

WO 2004/063359 PCT/EP2004/000099 332/357

aacgcgccag	aaacgccgtc	gaagccgtgt	gcgagacạcc	geggeegeeg	gcgttgtgga	9420
tacctcgcgg	aaaacttggc	cctcactgac	agatgagggg	cggacgttga	cacttgaggg	9480
gccgactcac	ccggcgcggc	gttgacagat	gaggggcagg	ctcgatttcg	gccggcgacg	9540
tggagctggc	cageetegea	aatcggcgaa	aacgcctgat	tttacgcgag	tttcccacag	9600
atgatgtgga	caagcctggg	gataagtgcc	ctgcggtatt	gacacttgag	gggcgcgact	9660
actgacagat	gaggggegeg	atccttgaca	cttgaggggc	agagtgctga	cagatgaggg	9720
gegeacetat	tgacatttga	ggggctgtcc	acaggcagaa	aatccagcat	ttgcaagggt	9780
ttccgcccgt	ttttcggcca	ccgctaacct	gtcttttaac	ctgcttttaa	accaatattt	9840
ataaaccttg	tttttaacca	gggctgcgcc	ctgtgcgcgt	gaccgcgcac	gccgaagggg	9900
ggtgececee	cttctcgaac	catacagges	cgctaacgcg	ggcctcccat	cccccaggg	9960
gctgcgcccc	tcggccgcga	acggcctcac	cccaaaaatg	gcagcgctgg	cagtccttgc	10020
cattgccggg	atcggggcag	taacgggatg	ggcgatcagc	ccgagcgcga	cgcccggaag	10080
cattgacgtg	ccgcaggtgc	tggcatcgac	attcagcgac	caggtgccgg	gcagtgaggg	10140
cggcggcctg	ggtggcggcc	tgcccttcac	ttcggccgtc	ggggcattca	cggacttcat	10200
ggcggggccg	gcaatttta	ccttgggcat	tcttggcata	gtggtcgcgg	gtgccgtgct	10260
cgtgttcggg	ggtgcgataa	acccagcgaa	ccatttgagg	tgataggtaa	gattataccg	10320
aggtatgaaa	acgagaattg	gacctttaca	gaattactct	atgaagcgcc	atatttaaaa	10380
agctaccaag	acgaagagga	tgaagaggat	gaggaggcag	attgccttga	atatattgac	10440
aatactgata	agataatata	tcttttatat	agaagatatc	gccgtatgta	aggatttcag	10500
ggggcaaggc	ataggcagcg	cgcttatcaa	tatatctata	gaatgggcaa	. agcataaaaa	10560
cttgcatgga	ctaatgcttg	aaacccagga	caataacctt	atagcttgta	aattctatca	10620

WO 2004/063359 PCT/EP2004/000099 333/357

taattgggta atgactccaa cttattgata gtgttttatg ttcagataat gcccgatgac 10680 tttgtcatgc agctccaccg attttgagaa cgacagcgac ttccgtccca gccgtgccag 10740 gtgctgcctc agattcaggt tatgccgctc aattcgctgc gtatatcgct tgctgattac 10800 gtgcagcttt cccttcaggc gggattcata cagcggccag ccatccgtca tccatatcac 10860 cacgtcaaag ggtgacagca ggctcataag acgccccagc gtcgccatag tgcgttcacc 10920 gaatacgtgc gcaacaaccg tcttccggag actgtcatac gcgtaaaaca gccagcgctg 10980 gcgcgattta gccccgacat agccccactg ttcgtccatt tccgcgcaga cgatgacgtc 11040 actgcccggc tgtatgcgcg aggttaccga ctgcggcctg agttttttaa gtgacgtaaa 11100 atcgtgttga ggccaacgcc cataatgcgg gctgttgccc ggcatccaac gccattcatg 11160 gccatatcaa tgattttctg gtgcgtaccg ggttgagaag cggtgtaagt gaactgcagt 11220 tgccatgttt tacggcagtg agagcagaga tagcgctgat gtccggcggt gcttttgccg 11280 ttacgcacca ccccgtcagt agctgaacag gagggacagc tgatagacac agaagccact 11340 ggagcacctc aaaaacacca tcatacacta aatcagtaag ttggcagcat cacccataat 11400 tgtggtttca aaatcggctc cgtcgatact atgttatacg ccaactttga aaacaacttt 11460 gaaaaagctg ttttctggta tttaaggttt tagaatgcaa ggaacagtga attggagttc 11520 gtcttgttat aattagcttc ttggggtatc tttaaatact gtagaaaaga ggaaggaaat 11580 aataaatggc taaaatgaga atatcaccgg aattgaaaaa actgatcgaa aaataccgct 11640 gcgtaaaaga tacggaagga atgtctcctg ctaaggtata taagctggtg ggagaaaatg 11700 aaaacctata tttaaaaatg acggacagcc ggtataaagg gaccacctat gatgtggaac 11760 gggaaaagga catgatgcta tggctggaag gaaagctgcc tgttccaaag gtcctgcact 11820 ttgaacggca tgatggctgg agcaatctgc tcatgagtga ggccgatggc gtcctttgct 11880

WO 2004/063359 PCT/EP2004/000099 334/357

cggaagagta tgaagatgaa caaagccctg aaaagattat cgagctgtat gcggagtgca 11940 tcaggctctt tcactccatc gacatatcgg attgtcccta tacgaatagc ttagacagcc 12000 gcttagccga attggattac ttactgaata acgatctggc cgatgtggat tgcgaaaact 12060 gggaagaaga cactccattt aaagatccgc gcgagctgta tgatttttta aagacggaaa 12120 agcccgaaga ggaacttgtc ttttcccacg gcgacctggg agacagcaac atctttgtga 12180 aagatggcaa agtaagtggc tttattgatc ttgggagaag cggcagggcg gacaagtggt 12240 atgacattgc cttctgcgtc cggtcgatca gggaggatat cggggaagaa cagtatgtcg 12300 agctattttt tgacttactg gggatcaagc ctgattggga gaaaataaaa tattatattt 12360 tactggatga attgttttag tacctagatg tggcgcaacg atgccggcga caagcaggag 12420 cgcaccgact tcttccgcat caagtgtttt ggctctcagg ccgaggccca cggcaagtat 12480 ttgggcaagg ggtcgctggt attcgtgcag ggcaagattc ggaataccaa gtacgagaag 12540 gacggccaga cggtctacgg gaccgacttc attgccgata aggtggatta tctggacacc 12600 aaggcaccag gcgggtcaaa tcaggaataa gggcacattg ccccggcgtg agtcggggca 12660 atcccgcaag gagggtgaat gaatcggacg tttgaccgga aggcatacag gcaagaactg 12720 atcgacgcgg ggttttccgc cgaggatgcc gaaaccatcg caagccgcac cgtcatgcgt 12780 gcgcccgcg aaaccttcca gtccgtcggc tcgatggtcc agcaagctac ggccaagatc 12840 gagcgcgaca gcgtgcaact ggctcccct gccctgcccg cgccatcggc cgccgtggag 12900 cgttcgcgtc gtctcgaaca ggaggcggca ggtttggcga agtcgatgac catcgacacg 12960 cgaggaacta tgacgaccaa gaagcgaaaa accgccggcg aggacctggc aaaacaggtc 13020 agcgaggcca agcaggccgc gttgctgaaa cacacgaagc agcagatcaa ggaaatgcag 13080 ctttccttgt tcgatattgc gccgtggccg gacacgatgc gagcgatgcc aaacgacacg 13140

WO 2004/063359 PCT/EP2004/000099 335/357

gcccgctctg	ccctgttcac	cacgcgcaac	aagaaaatcc	cgcgcgaggc	gctgcaaaac	13200
aaggtcattt	tccacgtcaa	caaggacgtg	aagatcacct	acaccggcgt	cgagctgcgg	13260
gccgacgatg	acgaactggt	gtggcagcag	gtgttggagt	acgcgaagcg	cacccctatc	13320
ggcgagccga	tcaccttcac	gttctacgag	ctttgccagg	acctgggctg	gtcgatcaat	13380
ggccggtatt	acacgaaggc	cgaggaatgc	ctgtcgcgcc	tacaggcgac	ggcgatgggc	13440
ttcacgtccg	accgcgttgg	gcacctggaa	teggtgtege	tgctgcaccg	cttccgcgtc	13500
ctggaccgtg	gcaagaaaac	gtcccgttgc	caggtcctga	tcgacgagga	aatcgtcgtg	13560
ctgtttgctg	gcgaccacta	cacgaaattc	atatgggaga	agtaccgcaa	gctgtcgccg	13620
acggcccgac	ggatgttcga	ctatttcagc	tcgcaccggg	agccgtaccc	gctcaagctg	13680
gaaaccttcc	gcctcatgtg	cggatcggat	tccacccgcg	tgaagaagtg	gcgcgagcag	13740
gtcggcgaag	cctgcgaaga	gttgcgaggc	agcggcctgg	tggaacacgc	ctgggtcaat	13800
gatgacctgg	tgcattgcaa	acgctagggc	cttgtggggt	cagttccggc	tgggggttca	13860
gcagccagcg	ctttactggc	atttcaggaa	caagcgggca	ctgctcgacg	cacttgcttc	13920
gctcagtatc	gctcgggacg	cacggcgcgc	tctacgaact	gccgataaac	agaggattaa	13980
aattgacaat	tgtgattaag	gctcagattc	gacggcttgg	agcggccgac	gtgcaggatt	14040
tccgcgagat	ccgattgtcg	gccctgaaga	aagctccaga	gatgttcggg	tccgtttacg	14100
agcacgagga	gaaaaagccc	atggaggcgt	tegetgaacg	gttgcgagat	gccgtggcat	14160
toggogocta	catcgacggc	gagatcattg	ggctgtcggt	cttcaaacag	gaggacggcc	14220
ccaaggacgc	tcacaaggcg	catctgtccg	gcgttttcgt	ggagcccgaa	cagegaggee	14280
gaggggtcgc	cggtatgctg	ctgcgggcgt	tgccggcggg	tttattgctc	gtgatgatcg	14340
tccgacagat	tccaacggga	atctggtgga	tgcgcatctt	catectegge	gcacttaata	14400

WO 2004/063359 PCT/EP2004/000099 336/357

tttcgctatt	ctggagcttg	ttgtttattt	cggtctaccg	cctgccgggc	ggggtcgcgg	14460
cgacggtagg	cgctgtgcag	ccgctgatgg	tcgtgttcat	ctctgccgct	ctgctaggta	14520
gcccgatacg	attgatggcg	gtcctggggg	ctatttgcgg	aactgcgggc	gtggcgctgt	14580
tggtgttgac	accaaacgca	gcgctagatc	ctgtcggcgt	cgcagcgggc	ctggcggggg	14640
cggtttccat	ggcgttcgga	accgtgctga	cccgcaagtg	gcaacctccc	gtgcctctgc	14700
tcacctttac	cgcctggcaa	ctggcggccg	gaggacttct	gctcgttcca	gtagctttag	14760
tgtttgatcc	gccaatcccg	atgcctacag	gaaccaatgt	teteggeetg	gcgtggctcg	14820
gcctgatcgg	agcgggttta	acctacttcc	tttggttccg	ggggatctcg	cgactcgaac	14880
ctacagttgt	ttccttactg	ggctttctca	gccccagatc	tggggtcgat	cagccgggga	14940
tgcatcaggc	cgacagtcgg	aacttcgggt	ccccgacctg	taccattcgg	tgagcaatgg	15000
ataggggagt	tgatatcgtc	aacgttcact	tctaaagaaa	tagcgccact	cagettecte	15060
agcggcttta	tccagcgatt	tcctattatg	teggeatagt	teteaagate	gacagcctgt	15120
cacggttaag	cgagaaatga	ataagaaggc	tgataattcg	gatetetgeg	agggagatga	15180
tatttgatca	caggcagcaa	cgctctgtca	tcgttacaat	caacatgcta	ccctccgcga	15240
gatcatccgt	gtttcaaacc	cggcagctta	gttgccgttc	ttccgaatag	catcggtaac	15300
atgagcaaag	tctgccgcct	tacaacggct	ctcccgctga	cgccgtcccg	gactgatggg	15360
ctgcctgtat	cgagtggtga	ttttgtgccg	agctgccggt	cggggagctg	ttggctggct	15420
ggtggcagga	tatattgtgg	tgtaaacaaa	ttgacgctta	gacaacttaa	taacacattg	15480
cggacgtttt	taatgtactg	gggtggtttt	tcttttcacc	agtgagacgg	gcaacagctg	15540
attgcccttc	accgcctggc	cctgagagag	ttgcagcaag	cggtccacgc	tggtttgccc	15600
cagcaggcga	aaatcctgtt	tgatggtggt	tccgaaatcg	gcaaaatccc	ttataaatca	15660

WO 2004/063359 PCT/EP2004/000099

aaagaatagc ccgagatagg gttgagtgt gttccagttt ggaacaagag tccactatta 15720
aagaacgtgg actccaacgt caaagggcga aaaaccgtct atcagggcga tggcccacta 15780
cgtgaaccat cacccaaatc aagtttttg gggtcgaggt gccgtaaagc actaaatcgg 15840
aaccctaaag ggagcccccg atttagagct tgacggggaa agccggcgaa cgtggcgaga 15900
aaggaaggga agaaagcgaa aggagcgggc gccattcagg ctgcgcaact gttgggaagg 15960
gcgatcggtg cgggctctt cgctattacg ccagctggcg aaagggggat gtgctgcaag 16020
gcgattaagt tgggtaacgc cagggttttc ccagtcacga cgttgtaaaa cgacggccag 16080
tgaattcgag ctcggtaccc ggg

<210> 63

<211> 25

<212> - DNA

<213> Artificial

<220>

<223> Primer

<400> 63

ggcgtacttg aaggaaccct taccg

<210> 64

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 64

attgatgete ceggteaceg tgatt

25

<210>	65				•			
<211>	500							
<212>	DNA							
<213>	<213> Blakeslea trispora							
<400>	65							
aatctat	caca	atgctccata	gactcacatt	gatattgtcg	aagatttcga	tgctgactta	60	
gtagago	caac	tacaaaagtt	agcagagaag	catgatttct	taatctttga	agaccgcaag	120	
tttgcag	gata	tcggtatgtg	aattctatct	atttttttc	tgatgtgtgc	atggatgact	180	
catgato	cata	ttcttaggta	atactgtcaa	gcatcaatat	ggcaagggcg	tttacaagat	240	
tgcttct	ttgg	tctcatatta	ctaatgctca	cacagttcct	ggagaaggta	ttatcaaggg	300	
acttgc	cgaa	gtcggcctcc	ctcttggtcg	tggcttgctt	ttgctagcag	aaatgtcatc	360	
tcaagg	tgca	ttaactaagg	gtatttacac	tgccgaatct	gtcaatatgg	ctcgccgcaa	420	
caaaga	tttc	gtttttggct	ttattgcaca	acacaaaatg	aatcagtatg	atgatgagga	480	
ttttgt	tgtc	atgtcgcctg					500	
ttttgt	tgtc	atgtcgcctg					500	
<210>	tgtc 66	atgtcgcctg					500	
<210>		atgtcgcctg					500	
<210>	66	atgtcgcctg					500	
<210> <211>	66 611 DNA	atgtcgcctg keslea trisp	pora				500	
<210> <211> <212> <213>	66 611 DNA Blal		pora				500	
<210> <211> <212> <213> <400>	66 611 DNA Blal	keslea trisp		aaaagaaatt	cagaagcata	gcacattott		
<210> <211> <212> <213> <400>	66 611 DNA Blal			aaaagaaatt	cggaagcatg	gcacattctt	500	
<210> <211> <212> <213> <400> gagatt	66 611 DNA Blal 66 aaaa	keslea trisp	aaagaaagtg					
<210> <211> <212> <213> <400> gagatt	66 611 DNA Blal 66 aaaa	keslea trisp tagataagga	aaagaaagtg gactttcttt	ttccatcgat	atgatatatg	catatgatag	60	
<210> <211> <212> <213> <400> gagatt cttttt	66 611 DNA Blal 66 aaaa ataa	keslea trisp tagataagga atacatgcct	aaagaaagtg gactttcttt aaggagtttg	ttccatcgat aaattttgtc	atgatatatg	catatgatag	60 120	
<210> <211> <212> <213> <400> gagatt cttttt atatac	66 611 DNA Blai 66 aaaa ataa ataa	keslea trisp tagataagga atacatgcct aatcttcttc	aaagaaagtg gactttcttt aaggagtttg acacaagaat	ttccatcgat aaattttgtc agttaccaat	atgatatatg ctccaggagc ttgctttggt	catatgatag aaaaaaaagt cttacgtgct	60 120 180	

WO 2004/063359 PCT/EP2004/000099 339/357

tttactaatt	cggcaaacat	ccaatttggc	aaacagcagc	ctgtgcaacg	ctctcgagat	420
gacagtatct	ttgattacac	tctaaatctc	gatgacccga	ccaaaaagag	cgaacaaaga	480
aataatcttg	tgcattcgaa	tatgatggaa	gatttttcc	cccttattct	aaatgttgac	540
atagcgtgta	tgttatataa	acaaaaagaa	attgtacaaa	ctttctttc	ttctcttttt	600
attttatctc	t					611

<210> 67

<211> 720

<212> DNA

<213> Blakeslea trispora

<400> 67

atgtcaatac tcacttatct ggaatttcat ctctactata cactacctgt ccttgcggca 60 ttgtgttggc tgctaaagcc gtttcactca cagcaagaca atctcaagta taaattttta 120 atgttgatgg ccgcctctac cgcatcgatt tgggacaatt atatcgttta tcatcgcgct 180 240 tggtggtact gtcctacttg tgttgtggct gtcattggct atgtacctct agaagaatac 300 atgttcttta tcatcatgac tttaatgact gtcgcgttct caaactttgt tatgcgttgg 360 cacttgcata ctttctttat tagacccaac acttcttgga agcaaacact attagtacgc cttgtgcctg tttcagcttt attggcaatc acttatcatg cttggcactt gacactgcca 420 aataaacctt cattttatgg ttcatgcatc ctttggtatg cttgtcctgt gttggctatt 480 ctttggctgg gtgctggcga atatatcttg cgtcgacctg tggctgtcct tttgtctatt 540 gttatcccta gtgtatacct atgttgggct gatatcgtcg ctattagtgc tggcacatgg 600 catatttctc ttagaacaag cactggcaaa atggtagtac ccgatttacc tgtagaagaa 660 720 tgcctgtttt ttactttgat caacacagtc ttggtttttg ctacctgtgc tatagaccgc

<210> 68

<211> 1089

<212> DNA

<213> Blakeslea trispora

<400> 68

ctgtacaaat	catctgttca	aaatcaaaac	cctaaacaag	ccatttccct	tttccagcat	60
gtcaaagagc	tagcatgggc	cttctgtctt	cctgaccaaa	tgctcaacaa	tgaattgttt	120
gatgatctta	ctatcagctg	ggatatttta	cgtaaagcct	caaagtcatt	ctatactgca	180
tctgccgttt	ttccaagtta	tgtacgtcaa	gacttgggtg	ttctctatgc	tttctgcaga	240
gctaccgatg	acctgtgcga	tgatgaatcc	aaatctgttc	aagaaagaag	agaccaatta	300
gatcttactc	gacaatttgt	tcgtgatctc	tttagccaaa	agaccagtgc	gcctattgtg	360
attgattggg	aattgtatca	aaaccaactt	cctgcttctt	gtatatcagc	ctttagagcc	420
tttactcgcc	ttcgccatgt	ccttgaagta	gaccctgtag	aagaactatt	agatggttac	480
aaatgggatc	ttgagcgtcg	tcctatcctt	gatgaacaag	acttggaggc	atactctgct	540
tgtgtggcca	gtagtgtggg	tgaaatgtgc	acacgtgtga	ttcttgctca	agaccaaaag	600
gaaaatgatg	cttggataat	tgaccgtgca	cgtgagatgg	ggctggtgct	acaatacgtt	660
aacattgctc	gagacattgt	gactgatagc	gagactctgg	gtcgatgtta	tctgcctcaa	720
caatggctta	gaaaagaaga	aacagaacaa	atacagcaag	gcaacgcccg	tagcctaggt	780
gatcaaagac	tgttgggctt	gtctctgaag	cttgtaggaa	aggcagacgc	tatcatggtg	840
agagctaaga	agggcattga	caagttgccg	gcaaactgtc	aaggcggtgt	acgagctgct	900
tgccaagtat	atgctgcaat	tggatctgta	ctcaagcagc	agaagacaac	atatcctaca	960
agagctcatc	taaaaggaag	cgaacgtgcc	aagattgctc	tgttgagtgt	atacaacctc	1020
tatcaatctg	aagacaagcc	tgtggctctc	cgtcaagcta	gaaagattaa	gagtttttt	1080

341/357

	4000
gttgattag	1089
<210> 69	
<211> 611	
<212> DNA	
<213> Blakeslea trispora	
<400> 69	
agagataaaa taaaaagaga agaaaagaaa gtttgtacaa tttctttttg tttatataac	60
atacacgcta tgtcaacatt tagaataagg gggaaaaaat cttccatcat attcgaatgc	120
acaagattat ttctttgttc gctctttttg gtcgggtcat cgagatttag agtgtaatca	180
	0.40
aagatactgt catctcgaga gcgttgcaca ggctgctgtt tgccaaattg gatgtttgcc	240
	300
gaattagtaa aatacgcaag catttettae ettteegete eetttteeta atteteecaa	300
Sponson	360
agactaaatg aggaaagata aaggacaaag aaaatgtaaa gacaaagaaa ttgaaaacga	
tataaacttg cagcacgtaa gaccaaagca aattggtaac tattcttgtg tacaaacatg	420
tataaacteg tageategeaa gatetaaagea aatteggeaan 1	
tataaaaaaa aactttttt tgctcctgga ggacaaaatt tcaaactcct tgaagaagat	480
•	
tgcttgtata tctatcatat gcatatatca tatcgatgga aaaagaaagt caggcatgta	540
tttataaaaa gaagaatgtg ccatgcttcc gaatttcttt tcactttctt ttccttatct	600
attttaatct c	611
<210> 70	
<211> 882	
<212> DNA	
<213> Haematococcus pluvialis	
<400> 70	
atgctgtcga agctgcagtc aatcagcgtc aaggcccgcc gcgttgaact agcccgcgac	60
40g01g10g1 4g00g11g10 111111g1g 11111 11 11 11 11 11 11 11 11	
atcacgcggc ccaaagtctg cctgcatgct cagcggtgct cgttagttcg gctgcgagtg	120

WO 2004/063359 PCT/EP2004/000099 342/357

gcagcaccac	agacagagga	ggcgctggga	accgtgcagg	ctgccggcgc	gggcgatgag	180
cacagcgccg	atgtagcact	ccagcagctt	gaccgggcta	tcgcagagcg	tegtgeeegg	240
cgcaaacggg	agcagctgtc	ataccaggct	gccgccattg	cagcatcaat	tggcgtgtca	300
ggcattgcca	tcttcgccac	ctacctgaga	tttgccatgc	acatgaccgt	gggcggcgca	360
gtgccatggg	gtgaagtggc	tggcactctc	ctcttggtgg	ttggtggcgc	gctcggcatg	420
gagatgtatg	cccgctatgc	acacaaagcc	atctggcatg	agtcgcctct	gggctggctg	480
ctgcacaaga	gccaccacac	acctcgcact	ggaccctttg	aagccaacga	cttgtttgca	540
atcatcaatg	gactgcccgc	catgetestg	tgtacctttg	gettetgget	gcccaacgtc	600
ctgggggcgg	cctgctttgg	agcggggctg	ggcatcacgc	tatacggcat	ggcatatatg	660
tttgtacacg	atggcctggt	gcacaggcgc	tttcccaccg	ggcccatcgc	tggcctgccc	720
tacatgaagc	gcctgacagt	ggcccaccag	ctacaccaca	gcggcaagta	cggtggcgcg	780
ccctggggta	tgttcttggg	tccacaggag	ctgcagcaca	ttccaggtgc	ggcggaggag	840
gtggagcgac	tggtcctgga	actggactgg	tccaagcggt	ag		882

<210> 71

<211> 528

<212> DNA

<213> Erwinia uredovora

<400> 71

atgttgtgga tttggaatgc cetgategtt ttegttaceg tgattggcat ggaagtgatt 60
getgeactgg cacacaaata catcatgcae ggetggggtt ggggatggca tettteacat 120
catgaacege gtaaaggtge gtttgaagtt aacgatettt atgeegtggt ttttgetgea 180
ttategatee tgetgattta tetgggeagt acaggaatgt ggeegeteea gtggattgge 240

WO 2004/063359 PCT/EP2004/000099 343/357

gcaggtatga	cggcgtatgg	attactctat	tttatggtgc	acgacgggct	ggtgcatcaa	300
cgttggccat	tccgctatat	tccacgcaag	ggctacctca	aacggttgta	tatggcgcac	360
cgtatgcatc	acgccgtcag	gggcaaagaa	ggttgtgttt	cttttggctt	cctctatgcg	420
ccgcccctgt	caaaacttca	ggcgacgctc	cgggaaagac	atggcgctag	agcgggcgct	480
gccagagatg	cgcagggcgg	ggaggatgag	cccgcatccg	ggaagtaa		528

<210> 72

<211> 762

<212> DNA

<213> Nostoc sp. PCC73102

<400> 72

60 atgatccagt tagaacaacc actcagtcat caagcaaaac tgactccagt actgagaagt 120 aaatctcagt ttaaggggct tttcattgct attgtcattg ttagcgcatg ggtcattagc 180 ctgagtttat tactttccct tgacatctca aagctaaaat tttggatgtt attgcctgtt atactatggc aaacattttt atatacggga ttatttatta catctcatga tgccatgcat 240 300 ggcgtagtat ttccccaaaa caccaagatt aatcatttga ttggaacatt gaccctatcc 360 ctttatggtc ttttaccata tcaaaaacta ttgaaaaaac attggttaca ccaccacaat ccagcaagct caatagaccc ggattttcac aatggtaaac accaaagttt ctttgcttgg tattttcatt ttatgaaagg ttactggagt tgggggcaaa taattgcgtt gactattatt 480 tataactttg ctaaatacat actccatatc ccaagtgata atctaactta cttttgggtg 540 600 ctaccctcgc ttttaagttc attacaatta ttctattttg gtactttttt accccatagt gaaccaatag ggggttatgt tcagcctcat tgtgcccaaa caattagccg tcctatttgg 660 720 tggtcattta tcacgtgcta tcattttggc taccacgagg aacatcacga atatcctcat 762 atttcttggt ggcagttacc agaaatttac aaagcaaaat ga

<210> 73 <211> 617 <212> DNA <213> Haematococcus pluvialis <400> 73 tagggtgcgg aaccaggcac gctggtttca cacctcatgc ctgtgataag gtgtggctag 60 agcgatgcgt gtgagacggg tatgtcacgg tcgactggtc tgatggccaa tggcatcggc 120 catgtctggt catcacgggc tggttgcctg ggtgaaggtg atgcacatca tcatgtgcgg 180 ttggaggggc tggcacagtg tgggctgaac tggagcagtt gtccaggctg gcgttgaatc 240 300 agtgagggtt tgtgattggc ggttgtgaag caatgactcc gcccatattc tatttgtggg agetgagatg atggeatget tgggatgtge atggateatg gtagtgeage aaactatatt 360 420 cacctagggc tgttggtagg atcaggtgag gccttgcaca ttgcatgatg tactcgtcat ggtgtgttgg tgagaggatg gatgtggatg gatgtgtatt ctcagacgta gaccttgact 480 540 ggaggcttga tcgagagagt gggccgtatt ctttgagagg ggaggctcgt gccagaaatg 600 gtgagtggat gactgtgacg ctgtacattg caggcaggtg agatgcactg tctcgattgt aaaatacatt cagatgc 617 <210> 74 <211> 1208 <212> DNA <213> Haematococcus pluvialis <400> 74 60 attgtgactg atagcgagac tctgggtcga tgttatctgc ctcaacaatg gcttagaaaa gaagaaacag aacaaataca gcaaggcaac gcccgtagcc taggtgatca aagactgttg 120

ggettgtete tgaagettgt aggaaaggea gaegetatea tggtgagage taagaaggge

180

WO 2004/063359 PCT/EP2004/000099 345/357

attgacaagt tgccggcaaa	ctgtcaaggc	ggtgtacgag	ctgcttgcca	agtatatgct	240
gcaattggat ctgtactcaa	gcagcagaag	acaacatatc	ctacaagagc	tcatctaaaa	300
ggaagcgaac gtgccaagat	tgctctgttg	agtgtataca	acctctatca	atctgaagac	360
aagcctgtgg ctctccgtca	agctagaaag	attaagagtt	tttttgttga	ttagtgaatt	420
tttgttttat ttatgtctga	tagttcaata	aagagacaac	acatacaata	taaaatcatt	480
gtctttaaat gttaatttag	tagagtgtaa	agcctgcatt	ttttttgtac	gcataaacaa	540
tgaattcacc ccgcttctgg	tttttaaata	attatgtcaa	actagggaaa	attcttttt	600
ttetettegt tettttttg	gcttgttgtg	gagtcacagg	cttgtcttca	gattgataga	660
ggttgtatac actcaacaga	gcaatcttgg	cacgttcgct	tccttttaga	tgagctcttg	720
taggatatgt tgtcttctgc	tgcttgagta	cagatccaat	tgcagcatat	acttggcaag	780
cagetegtae acegeettga	cagtttgccg	gcaacttgtc	aatgcccttc	ttagctctca	840
ccatgatagc gtctgccttt	cctacaagct	tcagagacaa	gcccaacagt	ctttgatcac	900
ctaggetacg ggcgttgcct	tgctgtattt	gttctgtttc	ttcttttcta	agccattgtt	960
gaggcagata acategacec	aacatcctcg	agccatacta	cagcataaaa	ggatacgttt	1020
tctttaacag aaatttaccc	ttttgttatc	agcacataca	aaaaaaaaga	aatttaagat	1080
gagtaggact tecattetet	caaaaatttt	attcaatcca	taaatgaatt	atttttggac	1140
aaaaaagaaa gattatgcct	gattttctct	atttttttt	tttttacaac	tccaccaata	1200
ctttctag					1208

<210> 75

<211> 6316

<212> DNA

<213> Blakeslea trispora

<220> <221> misc_feature <222> (2694)..(2694) <223> n is a, c, g, or t <220> <221> misc_feature <222> (4263)..(4263) <223> n is a, c, g, or t <400> 75 aaggatgaag aatccaactc taataaaaat cttatggata tctttgatcg actcaaaaag 60 gctttcaatg ctattgctat taaaaaaaaa gagagagaga gaactatgag caaaaggact 180 ctatgccaag atggcaaaaa ggcaccagaa accettagtt tattattgca taatccagtc gagctagtac ttctgtagct caagcttaac cgaggatctt ggaatcaact cgtctcgtca 240 ctcttgccga tgatcctaga aatggtatct atggatgtta tactaacatt gttatctttc 300 360 aaggeetega agatgttatt gttgeggtga taaatagget getatgtaet gaagttgete 420 tgtaaaatga atctagttca ctgcctactc agcaaatggt tgtttctaat gtctttaaag aaagaaaaaa agatacatat agactaccct tcctttcaag actgtaatcg agaatcggcc 480 gatggtttat tacaattaga cgctgggaat aagcaaaagg attcatcttt gtaaataaga 540 600 gactggtgca tatgaaagca aggatcgtat caaggaatag ttttgatcga gcatcaccag caaatgctgc taatgttggc ttcttctttg cttcctgaga ttgaatggga tgtgcctaga 660 gcattgctat ttttaagtgt atactttaga tttgtgtctt tagatttgtg tcattttatt 720 780 tagtcaagaa agatccccct ttctctatgt atgctaagaa gaaggagcaa gaagtgtatt tacaagttgg aatgagattg aaatattgta cataataata ataaaaagaa aggtagatca 840

aaaaaaatgt tctgcctatt gtaagaaatc gggaccaaca ggtgcttgat aaccagaagt

900

WO 2004/063359 PCT/EP2004/000099 347/357

agcttccaat	tcaggtagag	gctctaggga	caaatacaca	attatgacag	gaattttctt	960
gttgacttga	acactacaag	agaaacgggt	cagcacaaaa	tccgaaaaaa	aaaagaaacg	1020
gaccattcat	gtcttaccta	tctagctctt	tgtcttcaat	tgcatcccat	tgctcaacca	1080
cagatacgct	tcccaattga	gtatattgat	gaagtgttcc	ctgcattttt	cgcttgacta .	1140
attccactac	agtcacagtc	ttattaatgt	tttgtccttt	accagtcagg	ataatatgat	1200
ctttttgctt	cttctatcaa	aaaaataatt	cttgttttga	ataaaaaaaa	caaatattta	1260
aagaaactac	tttgatgacg	gtacctggaa	taactcgaga	cacacatcta	catatgcgtt	1320
gattttattg	tggctaattc	gaacctcatt	ttctgctggt	gggggctgtt	gactttcagt	1380
tgctgagacg	tecttettge	ttcttttata	gtcttccact	atgattttaa	tcaagaaagt	1440
aagtcagtga	tgattgttac	aagctatata	tcttgaaaaa	gaacagagag	gtattattat	1500
cagatgcaac	atggttttct	gtatcatttt	catttcagtt	tctctgttca	aaaaaaaaaa	1560
gaacactttc	tctttccact	cctcaaattt	tttctgctaa	actcctcgca	aaacatgtat	1620
ttgctttaaa	ctacaagttg	caattgtctg	atttagcaat	ttcaatatgc	cttttgtgaa	1680
tccacccaaa	aataaacaag	tgcttgagta	tacttgggtt	cagttcaaaa	gaaagcaagc	1740
tttttttt	ctttcttggg	aaagaaaaaa	aaatattgtt	gagccatcct	ttaccagcag	1800
tatgcgagct	acgacatagc	tggtctaaca	atgactgcaa	gcaatagatc	gagcttagtc	1860
tttctattgc	ttcyttgttt	gatctatgtt	cggccttacg	ctgacctatc	caatactcga	1920
gataggcaac	aagatttcga	acagtaatga	aataaatttc	ggataacagt	tgtggatgag	1980
gaagagaaag	cgacttgaac	tcgagaaact	ttgttgaaat	gaaatccgac	cttttacgtg	2040
atcatcatgt	attatcctct	ttttctttt	tttcgtagtg	aattacttac	tgattgcgct	2100
caagtegegt	ctttataaag	aagaaaaaaa	aatattagaa	ctttcaaaaa	atataactga	2160

WO 2004/063359 PCT/EP2004/000099 348/357

aaataaaagt gtggctcgg	a gagcaaatac	cacatccttt	gtcttcgctt	tggtaacacg	2220
gttaataagc cactatagg	t gaataatgat	catttctgag	aataaagcgc	ggcttgaagc	2280
ttatatccat atcaggatt	c atattaggca	caactcacaa	ttgaggttcc	agaagtgcca	2340
atttttttt cctgatagc	c tgtccaatta	agatcaaaaa	ccactgagtt	ttctctatat	2400
atttttttt ttcataatt	c ttaactcttc	ttcctctctc	tetetetete	tctctttttg	2460
gcttgcaaaa aaaatcttt	a gtaataccaa	agaaagcaaa	ccttttcctt	ttcttatttc	2520
cttgcttgtt ttttaattt	t tgatttctct	atgctttaaa	tacccatttc	tttctttctt	2580
ctgctattac ctatctttt	c attectetee	ccctctctc	tcttggtcta	taaacatcat	2640
gaagteetet tttaaaagt	t cgcttgacat	ttatgctgtt	tatatacagc	atcntgtgtt	2700
ttccaagtgg ttcattctt	g cttttgttct	ttcgattttc	ctcaacactt	atctactgaa	2760
cgcttcgaag caacagccc	a aagtgataat	caaaaaggtt	attgagcggg	tagaagtacc	2820
aagtagagaa caacctaaa	t cagtcataaa	gecetectee	aagaaacact	cttctcatca	2880
tcagtctgat gtcattcgc	c ctcttgatga	agtattgggt	ttgctcggaa	cacccgaggc	2940
cttgactgat gaagagatc	a tctctattgt	tcaagctggt	aaaatggccc	cctatgctct	3000
tgaaaaggtc ttgggcgat	t tagagcgcgc	tgtccatatc	cgtcgtgctt	tgatctcccg	3060
tgactctcgt acgaaaact	t tggaagacag	tatgcttccc	gtgaaaaact	atcattatga	3120
taaagtcatg ggtgcttgt	t gtgaaaatgt	cattggttat	atgcctattc	cagtaggtgt	3180
cgcaggtaag aagttcaac	a agtcgcgata	tttgacaagt	tgctcatcat	tttcgaaaca	3240
ggtcctttgg tgattgatg	g tgattctatt	catattccca	tggcaactac	ggaaggttgt	3300
ttagttgctt ctactgcca	g aggttgtaaa	gcaatcaatg	ctggtggtgg	tgccaacaca	3360
attgttgttg ctgatggta	t gactcgaggt	ccttgtgtcg	aatttcctac	aatcactcgc	3420

WO 2004/063359 PCT/EP2004/000099 349/357

gctgctgact	gtaaacgatg	gattgaacaa	gagggtgaag	ctatcgtgac	cgaggcattc	3480
aattcaactt	ctcgttttgc	tcgtgttcgt	aaattgaaag	ttgctcttgc	cggtcgtcta	3540
gtctacatcc	gtttctctac	cactacaggt	gatgcaatgg	gcatgaacat	gatctccaag	3600
ggttgtgaaa	aggctttaag	caagattgct	gagagatatc	ctgatatgca	gatcatttct	3660
ctttctggta	actattgtac	tgacaagaaa	cctgctgcta	tcaactggat	tgaaggacgt	3720
ggtaaatctg	ttgttgctga	sgctgtcatc	cctggtacgg	ttgtcgaaaa	ggtattgaag	3780
acctctgtta	gtgctttggt	tgagctgaac	atctctaaaa	acctggttgg	ttctgctatg	3840
gctggctccg	teggtggett	taacgctcat	gctgctaata	ttctaactgc	catttacctt	3900
gctactggtc	aagatcctgc	tcaaaatgta	sagagttcta	actgtattac	tttgatgaaa	3960
gctgtcaatg	gcgaaagaga	ccttcatatc	tcttgtacaa	tgccctgtat	tgaagtaggc	4020
accattggtg	gtggtactat	tttgcctcct	caacaagcca	tgttggattt	cattggtgtg	4080
cgtggtcctc	accctaccga	acctggtgcc	aatgcccgwc	gccttgctcg	tgttatctgt	4140
gcctctgtga	tggctggtga	attgtcttta	tgtgcagctt	tggctgctgg	tcatcttgta	4200
aaggcacaca	tggctcataa	tcgtaatacc	actgctgctg	ccgctgttgt	tectgecect	4260
aanggcatag	ttgatgtctc	tacacctcct	gctacacctg	cagaaaagaa	tgatcctatt	4320
cctggaagtt	gtatcaagtc	atagaattaa	tattatatat	atatcatata	caaaaaaaag	4380
аааааааааа	cactacatct	atttatattt	ctccatgtac	acacacacac	acacatataa	4440
aaactcttta	ttttccaata	ttttgctttt	ataaataatc	ttatttcatt	ctaaataaac	4500
tgttttttt	tattaatcat	caaaccctgc	tgagagctgt	gcaatatcat	ctatgttttc	4560
atggtttaac	tctggtatcg	gwcgagcctc	ctctgtactt	gaagtttgta	ggcagttttt	4620
atttaaggct	gctggtcgat	catgatcatc	akcaaacctg	acagcatgaa	gttttgactg	4680

atgagcaatt	tcactaaggg	cagaatctga	actctttcgc	ttcctactat	tgaccatatt	4740
gtctttaggt	ggaatgagtg	aatagcgtct	tgtcatatgt	aacacagaat	caacaatatc	4800
ctggtgatga	aactcggcca	aacatagcgc	ctttctcccc	caacaattat	aataatcaaa	4860
atgagaatga	catgtacggt	tttcctcgat	gacaatatcc	aacgtcttgt	cataatcctc	4920
tgtgcgyata	ccattcatct	tttggaagaa	cgcacggtag	ctctcacaag	ctgtcctcag	4980
agagttccgt	gccatgtttc	ccaatgctcc	tggcaagtcg	aaatgaagtt	gtcgaatctg	5040
gcgatgtatg	tctacaatgt	cgcctgtttc	tttcattaga	tcaagcattc	gtgtagccca	5100
aatgatgtct	atgttatgat	tttctttcat	tccagtaata	actatagttt	ctcggcaaat	5160
cgaatgastg	atggagtaaa	ttcatcaaaa	gtgcaagtaa	tacatacagt	gcttgaagaa	5220
atcttgtgta	gcacgcctat	attatgtaat	ataggatcga	ttctcgaaac	tcgacataac	5280
caccaggctt	tagcaagcgt	tttatttcat	tcatgacaag	ctattgttaa	ttcytgctta	5340
ataaaacaaa	atgaaaaaaa	cataccccc	tcmaaactta	cttcccactc	ttgattggaa	5400
aaacaggtat	agacgtgacg	catatgtata	taatcaaaac	actcatcagg	atagggtaaa	5460
ccattgagca	catcgcattg	ggtgaagaaa	gtattaggag	gcttgatggc	tgtaggatat	5520
ataggtgcaa	tatcaatacc	gtaaaactca	gcatttggga	attetgtage	catctccaga	5580
atccaagtac	ctgtgccaca	agcaacatca	agcactttag	gtaagggtat	acattgttgt	5640
tcttgttgtt	gttgttgaca	atcacttgag	tctgagtttc	gttttgattg	ttttaatgac	5700
aataattctt	ttacaggtgc	tgagaaatta	ccgtcaaata	gatacttgta	aataaaatgc	5760
taaaaataaa	aacaatagaa	aaaaaaattg	acgctcattt	cattactatg	gaaataactg	5820
caaaatctta	ı ccacttgtac	aagtctatct	tgctcaatct	catcgtttgg	cagaatgtat	5880
ttattgttgt	agtattgata	tcttctacca	ttcatgatat	aactgtcgct	tctaatgctc	5940

WO 2004/063359 PCT/EP2004/000099 351/357

tgaggtgaag	tacttgtagg	tgaaggtgga	agtgacgcaa	ttttgtcaag	cttaacagga	6000
tectetegge	tacatgtttt	ctgcatatca	ggaaaatctt	gtttatttga	aacatcaaca	6060
gtagatgtgg	tgtgatcttt	tttgaaaata	tcgatgcctt	cctttgaaag	ccttttgaaa	6120
ggctctttta	acttttttga	gtgagagcta	cccatgatag	cttatgaaga	attaaaaaga	6180
aaaaagcaaa	aaaaattaaa	aaaaaaaaaa	gtagcaaaaa	attctgtcgt	aattatacaa	6240
gccaatcaaa	atcgaaattc	atgcaaggca	tagatgttca	cgtggatttg	atggttgatc	6300
ctttttttt	gcaaga					6316

<210> 76

<211> 1170

<212> DNA

<213> Thermus thermophilus

<400> 76

atgaagcgcc tttccctgag ggaggcctgg ccctacctga aagacctcca gcaagatccc 60 120 ctegecgtee tgetggegtg gggeegggee cacecegge tetteettee eetgeeeege 180 🔨 ttccccctgg ccctgatctt tgaccccgag ggggtggagg gggcgctcct cgccgagggg 240 accaccaagg ccaccttcca gtaccgggcc ctctcccgcc tcacggggag gggcctcctc accgactggg gggaaagctg gaaggaggcg cgcaaggccc tcaaagaccc cttcctgccg 300 360 aagaacgtcc gcggctaccg ggaggccatg gaggaggagg cccgggcctt cttcggggag tggcgggggg aggagcggga cctggaccac gagatgctcg ccctctccct gcgcctcctc gggcgggccc tcttcgggaa gcccctctcc ccaagcctcg cggagcacgc ccttaaggcc 480 540 ctggaccgga tcatggccca gaccaggagc cccctggccc tcctggacct ggccgccgaa gcccgcttcc ggaaggaccg gggggccctc taccgcgagg cggaagccct catcgtccac 600

WO 2004/063359 PCT/EP2004/000099 352/357

ccgcccctct cccaccttcc ccgagagege gccc	tgagcg aggccgtgac ceteetggtg 660	
gegggceacg agaeggtgge gagegeeete acet	ggteet tteteeteet eteceacege 720	
ccggactggc agaagcgggt ggccgagagc gagg	aggegg ecetegeege etteeaggag 780	
gecetgagge tetacecece egectggate eteac	eccgga ggctggaaag gceceteete 840	
ctgggagagg accggctccc cccgggcacc accc	tggtcc tctccccta cgtgacccag 900	
aggetecact teecegatgg ggaggeette egge	ccgagc gcttcctgga ggaaaggggg 960	
accepttegg ggegetaett cecetttgge etgg	ggcaga ggctctgcct ggggcgggac 1020	
ttcgccctcc tcgagggccc catcgtcctc aggg	cettet teegeegett eegeetagae 1080	
cccctcccct tcccccgggt cctcgcccag gtcac	ecctga ggcecgaagg cgggetteec 1140	
gegeggeeta gggaggaggt gegggegtga	1170	

<210> 77

<400> 77

60 tctagaattc attccattcg aaaggatcaa cataaccaat ttaatgacta ctagctaatg 120 gatacaaata tacgcacaaa aaaagaaaga attctatgat caaagagaac acagacacag agtgatacat ttaaatggtt aagttcttat gatgttaaaa tggtaacttt attattgaat 180 240 taaatgcgaa tatcgttgct gctttgtact tggaaaacgt taggtaaaag ttggttaatg 300 aaagaagcag gagttgtagt atcatctctt gggaagaaat agaaaaagag gaaagtaaca aagtaacaag caagacaata atagatccaa tggctttcgg tcttacgagt ttgttcagga 360 gcatacttct tttggctatc ttgtaacttt cttggtaagg gattctggcc aaagctttta 420 480 cagacttggt cggaagtaag cttacttcca gcaagaacga taggaacacc agtacctgga

<211> 2981

<212> DNA

<213> Blakeslea trispora

WO 2004/063359 PCT/EP2004/000099 353/357

tgtgtactac	aaagaaaaga	gaaatgagta	cgtgcgttat	taaaaaaaag	aaaaaaagag	540
ggcaaaagta	ttacctagct	ccgacaaaga	aaagattatc	ataacggttt	gtggaatcct	600
tggtactagg	tctgaaccag	agaacttgga	acacatcatg	agaaagacca	agaatagaac	660
ctctccaaag	gttaaacttg	ctttgccaaa	cactaggatc	attcacttct	tcatgttcaa	720
tcaaattagc	aaagttgttt	actcccaaac	gacgttcgat	aacttccaga	accatcttgc	780
gtgcacggtt	taccaactca	ggataatttt	cttcagcact	gtttcctgtc	ttactcttca	840
tatggccaat	tggaaccaac	acaataatgg	agtccttgtt	gggaggtgcg	gcagattcat	900
caattcgaga	tggaacgttg	acatagaatg	aagcttcaga	gggcaaaccg	aagtcgttga	960
aaatctcatc	aaaactttcc	ttgtaggctt	cagccaagaa	gatattgtgt	acgtctaatt	1020
gaggcacctt	tgttgacatg	gaccaataaa	acgaaataga	tgatgaagtg	agtttctttg	1080
aggctaatgt	cttctttgtc	caattgcaag	gaggtaacag	atggtgataa	gcataaacaa	1140
gatccgcatt	acatacgact	gcatcggctt	caatgacttc	tccgctttcc	aaagtgacac	1200
cggttacacg	cttgtcttta	tcgacagtgt	taattttagc	aacaggcgat	tgatatctga	1260
attcagcacc	gtactttttg	gaggcgatag	actcaagctt	ctgaacaacc	atgttgaaac	1320
caccacgagg	ataccagata	ccttcagcaa	actcggtgta	ttgtaacaaa	ctgtaaactg	1380
ctggagcatc	ataaggcgac	atactatatt	ccaaaaatag	aaaatagaac	aatgaatatc	1440
aaaattcctt	tcacttgccc	tttttcacat	ttctcttttc	ccacccccga	ccggtctcac	1500
tcatttttt	ttcatcccac	accacgcgtt	gtatgtgtac	ttaccccata	tacattgttt	1560
gaaaagtaaa	agccatacgc	attttcttgg	tttggaaata	tttactggct	cggtcataga	1620
tcttaccaaa	caagtgcaag	cgaaagattt	caggcacata	ctgaagacga	atcaaatccc	1680
aaatggtttc	aaagttgcgc	ttgatagcaa	taaatgtacc	ttgttcataa	tggacatgtg	1740

WO 2004/063359 PCT/EP2004/000099 354/357

					•	
tttccttcat	gaaatccaag	aatctaccaa	atccaagggg	accctcaata	cggtccaatt	1800
cgcccttcat	cttggttaaa	tcggaagaga	gttgtacggc	atcaccgtcg	tcaaaatgaa	1860
ccttatagtt	attgtcacag	cgaagcaaat	ccaaatgatc	accaatacgt	tcatccaaạt	1920
cagcaaatgc	atcttcaaaa	agcttaggca	tcaaatagag	tgagggaccc	tgatcaaagc	1980
gatgaccatc	gtgatgaatg	aatgaacaac	ggccaccgga	aaagtcgttc	ttttcaacaa	2040
cagtaactcg	aaaaccttca	cgagcaagac	gagcagcagt	agcagttccg	ccaataccgg	2100
caccaatgac	aacaatatgc	ttcttttgat	cagacatgag	attaaaatag	ataaggaaaa	2160
gaaagtgaaa	agaaattcgg	aagcatggca	cattcttctt	tttataaata	catgcctgac	2220
tttcttttc	catcgatatg	atatatgcat	atgatagata	tacaagcaat	cttcttcaag	2280
gagtttgaaa	ttttgtcctc	caggagcaaa	aaaaagtttt	tttttataca	tgtttgtaca	2340
caagaatagt	taccaatttg	ctttggtctt	acgtgctgca	agtttatatc	gttttcaatt	2400
tetttgtett	tacattttct	ttgtccttta	tctttcctca	tttagtcttt	gggagaatta	2460
ggaaaaggga	gcggaaaggt	aagaaatgct	tgcgtatttt	actaattcgg	caaacatcca	2520
atttggcaaa	cagcagcctg	tgcaacgctc	tcgagatgac	agtatctttg	attacactct	2580
aaatctcgat	gacccgacca	aaaagagcga	acaaagaaat	aatcttgtgc	attcgaatat	2640
gatggaagat	tttttcccc	ttattctaaa	tgttgacata	gcgtgtatgt	tatataaaca	2700
aaaagaaatt	gtacaaactt	tetttette	tctttttatt	ttatctctat	gtcaatactc	2760
acttatctgg	aatttcatct	ctactataca	ctacctgtcc	ttgcggcatt	gtgttggctg	2820
ctaaagccgt	ttcactcaca	gcaagacaat	ctcaagtata	aatttttaat	gttgatggcc	2880
gcctctaccg	catcgatttg	ggacaattat	atcgtttatc	atcgcgcttg	gtggtactgt	2940
cctacttgtg	ttgtggctgt	cattggctat	gtacctctag	a		2981

<210> 78

<211> 1749

<212> DNA

<213> Blakeslea trispora

<400> 78

atgtctgatc	aaaagaagca	tattgttgtc	attggtgccg	gtattggcgg	aactgctact.	60
getgetegte	ttgctcgtga	aggttttcga	gttactgttg	ttgaaaagaa	cgacttttcc	120
ggtggccgtt	gttcattcat	tcatcacgat	ggtcatcgct	ttgatcaggg	tccctcactc	180
tatttgatgc	ctaagctttt	tgaagatgca	tttgctgatt	tggatgaacg	tattggtgat	240
catttggatt	tgcttcgctg	tgacaataac	tataaggttc	attttgacga	cggtgatgcc	300
gtacaactet	cttccgattt	aaccaagatg	aagggcgaat	tggaccgtat	tgagggtccc	360
cttggatttg	gtagattctt	ggatttcatg	aaggaaacac	atgtccatta	tgaacaaggt	420
acatttattg	ctatcaagcg	caactttgaa	accatttggg	atttgattcg	tcttcagtat	480
gtgcctgaaa	tctttcgctt	gcacttgttt	ggtaagatct	atgaccgagc	cagtaaatat	540
ttccaaacca	agaaaatgcg	tatggctttt	acttttcaaa	caatgtatat	gggtatgtcg	600
ccttatgatg	ctccagcagt	ttacagtttg	ttacaataca	ccgagtttgc	tgaaggtatc	660
tggtatcctc	gtggtggttt	caacatggtt	gttcagaagc	ttgagtctat	cgcctccaaa	720
aagtacggtg	ctgaattcag	atatcaatcg	cctgttgcta	aaattaacac	tgtcgataaa	780
gacaagcgtg	taaccggtgt	cactttggaa	agcggagaag	tcattgaagc	cgatgcagtc	840
gtatgtaatg	cggatcttgt	ttatgcttat	caccatctgt	tacctccttg	caattggaca	900
aagaagacat	tagcctcaaa	gaaactcact	tcatcatcta	tttcgtttta	ttggtccatg	960
tcaacaaagg	tgcctcaatt	agacgtacac	aatatcttct	tggctgaagc	ctacaaggaa	1020

WO 2004/063359 PCT/EP2004/000099 356/357 tttgatg agattttcaa cgacttcggt ttgccctctg aagcttcatt ctatgtcaac 1080

	agttttgatg	agattttcaa	cgacttcggt	ttgccctctg	aagcttcatt	ctatgtcaac	1080
	gttccatctc	gaattgatga	atctgccgca	cctcccaaca	aggactccat	tattgtgttg	1140
	gttccaattg	gccatatgaa	gagtaagaca	ggaaacagtg	ctgaagaaaa	ttatcctgag	1200
	ttggtaaacc	gtgcacgcaa	gatggttctg	gaagttatcg	aacgtcgttt	gggagtaaac	1260
	aactttgcta	atttgattga	acatgaagaa	gtgaatgatc	ctagtgtttg	gcaaagcaag	1320
	tttaaccttt	ggagaggttc	tattcttggt	ctttctcatg	atgtgttcca	agttctctgg	1380
	ttcagaccta	gtaccaagga	ttccacaaac	cgttatgata	atcttttctt	tgtcggagct	1440
	agtacacatc	caggtactgg	tgttcctatc	gttcttgctg	gaagtaagct	tacttccgac	1500
,	caagtctgta	aaagctttgg	ccagaatccc	ttaccaagaa	agttacaaga	tagccaaaag	1560
i	aagtatgctc	ctgaacaaac	tcgtaagacc	gaaagccatt	ggatctatta	ttgtcttgct	1620
:	tgttactttg	ttactttcct	ctttttctat	ttcttcccaa	gagatgatac	tacaactcct	1680
ţ	gettetttea	ttaaccaact	tttacctaac	gttttccaag	tacaaagcag	caacgatatt	1740
(cgcatttaa						1749

<210> 79

<211> 25

<212> DNA

<213> Artificial

<220>

<223> Primer

<400> 79

ccgatggcga cgacggaagg ttgtt

25

<210> 80

<211> 25

<212> DNA

WO 2004/063359

PCT/EP2004/000099 357/357

<213> Artificial

<220>

<223> Primer

<400> 80

catgttcatg cccattgcat cacct

25